University of Alberta

Trapped Between Paradigms: Composition Pedagogy in the Context of a Twelfth Grade Standardized Writing Assessment

by

David Hendrik Slomp ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Secondary Education

Edmonton, Alberta
Fall, 2007

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-33069-2
Our file Notre référence
ISBN: 978-0-494-33069-2

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

For Christian and Kiara:

May you too follow your dreams.

ABSTRACT

North American education is currently undergoing what some might call a revolution driven by test-based accountability (Brennan, 2006). History has taught us that some revolutions engender improvements while others bring little more than grief and suffering. The question that looms behind this current revolution—a question too often left unasked—is, "What effect is test-based accountability exerting on education"?

The goal of this study is to engage with this question within the context of Alberta's English 30-1 (academic grade 12 English) program. It focuses on understanding what effect the writing component of Alberta's English 30-1 exam is exerting on composition pedagogy and learning in Alberta's English 30-1 classes.

This study progresses in three phases. Phase one examines the literature on composition pedagogy to determine what understandings have emerged at the forefront of the field. Phase two examines Alberta's English 30-1 writing exam with a view to defining the construct it is measuring. These two sets of understandings are then compared with one another. Conclusions drawn from this analysis suggest that the English 30-1 diploma exam is based upon a theory of writing that has been discredited for more than three decades now: Essentially, this exam is designed to measure polished first draft writing.

The third phase of this study explores the impact that these flaws in assessment design might be exerting on teaching and learning in English 30-1 classes. Data was collected through a mixed methods—case studies and survey—design. Findings from this phase of the study suggest that the diploma exam supports a limited form of composition pedagogy. This study concludes that for students this exam encourages the development of skills related to creating polished first drafts rather than skills needed to successfully manage a robust multi-draft process. Together these findings suggest that Alberta's English 30-1 diploma exam is working counter to the goal of improving educational quality.

This study calls for reform of the Alberta's English 30-1 diploma exam. It recommends that current methods of reporting student results be reconsidered, and it suggests changes to school board policies related to the use of diploma exam results.

ACKNOWLEDGEMENTS

I would like to thank the following people for their support:

Anne, Brian, and Heather for inviting me into your classrooms and your lives;

Ingrid Johnston and Jacqueline Leighton for your kindness and your advice;

Margaret Iveson for providing me with a model of a graduate student supervisor that I will endeavor to emulate throughout my academic career, and for your kindness, wisdom, and friendship;

Betsy Sargent, Jill McClay, and Kathleen Blake Yancey for your supportive feedback;

James Nahachewsky for helping me keep this dissertation and this PhD endeavor in perspective;

The Department of Secondary Education, the Killam Trust, and SSHRC for provided me with the financial means to make this degree possible;

My parents for their ongoing love and support;

My wife Loretta for enabling me to chase my dreams, and for your constant love and encouragement;

My Heavenly Father who strengthens and guides me.

TABLE OF CONTENTS

Chapter 1: Foundational Perspectives 1

Writing My Life: Selected Episodes of My Personal Journey 2

Steps on the Journey 3

Grade Eleven: Directions 3

Grade 12: Writing for the Provincial Exam 4

Undergraduate Choices 6

Beginning Teacher, Finding my Way 8

Teaching Writing and the Provincial Writing Assessments 13

PhD Studies: Refining the Question 16

Epistemological and Ontological Issues 20

Focus on Validity 29

Research as Transaction 29

Implications for Research Design 34

Chapter 2: Accountability, Validity, and Composition Pedagogy 36

The Question of Ethics in Test-based Accountability in Canadian Education 37

Why has the Lapse Occurred? 40

The Ethics of Testing 43

The Ethics of Consequential Validity 46

Consequential Validity Within a Renewed Validity Framework 50

Challenges to Messick's Validity Theory 55

Over Emphasis on the Developer 56

Emphasis on Large-Scale Assessment of Consequential Validity 57

Emphasis Favors Test Score use Rather than Test Development 59

Consequential Validity, an Overreaction to Legal Climate 61

Further Challenges to Messick 65

Consequential Validity: What is the Evidence? 68

Historical Issues in Writing Assessment 73

Standardized Testing and the Teaching of Writing 79

Growth in Research on Writing Processes 80

Writers on Writing 88

Motivation 89

Processes 90

The First Stage 90

The Second Stage 93

The Post-process Movement 96

Regimentation of Process Pedagogy 97

Failure to Teach Necessary Skills and Conventions 99

The Writer Without Context 100

Implications for Teaching and Testing 102

Validity, a Pedagogical Issue 111

Chapter 3: Alberta's English 30-1 Diploma Exam 113

Construct-based Challenges to Writing Assessments 113

The Exam in Context 116

Writing Requirements in the English 30-1 Program of Studies 117

The English 30-1 Writing Exam 118

Personal Response to Text Assignment 119

Critical/Analytical Response to Literary Texts Assignment 121

Summary: The Test and the Construct 123

Comparing Constructs: The Test and the Theory 125

Construct Underrepresentation and Irrelevant Variance 126

Chapter 4: Mixed Methods Research 130

Mixed Methods Research 132

Advantages of Mixed Methods Research 134

Drawing on the Strengths of Quantitative and

Qualitative Designs 136

Limitations of Mixed Methods Research 137

Mixed Data Collection Methods 138

Chapter 5: Three Teachers 140

Case Study Methods 140

Selecting Participants 142

Data Collection 143

Teacher Interviews 144

Interview Data Analysis 145

Student Interviews 148

Interview Data Analysis 149

Classroom Observations 149

Analysis of Observation Data 150

Document Analysis 152

Analysis of Documents 152

Case Study Profiles 153

Anne's Profile 154

Perspectives on the Diploma Exam 174

Brian's Profile 177

Descriptions of Practice 181

Perspectives on the Diploma Exam 193

Positive Attributes 193

Critical Perspectives 197

Heather's Profile 200

Analysis Across Cases 223

Sources of Pedagogy 223

Senior High ELA Program of Studies 224

Previous Educational Experiences 224

Teacher Education Program 225

Diploma Exam Experiences 229

Narrowing Writing Focus 230

Narrowing the Marking Scheme 233

Limiting Process 234

Context and Experience 235

Attitudes Toward the Exam 235

Questioning the Exam's Construct 236

External Pressures 237

Teaching Experience 239

Caught Between Paradigms 240

Chapter 6: Twelfth Grade Student Writers 243

Methods and Analysis 243

Classroom Experiences of Survey Respondents 246

What Students Value in Their Own Writing 248

Student Analysis of Classroom Writing Assessments 249

        Student's Interpretation of the Exam's Construct 253
        Writing Process for Classroom Assessments 257
        Diploma Exam Writing Process 264
    What is the Impact? 266

Chapter 7: Conclusions and Implications 267
    Questions Asked: 268
    Questions Answered: 269
        A Limited Theory of Writing 269
        Negative Impact on Composition Pedagogy 269
        Poor Reporting Practices 270
        External Pressure Drives Teaching to the Test 271
        Critical Awareness Influences Pedagogy 272
        Inflation of Test Scores not a Primary Motivation 272
        Diploma Exam Supports Limited Learning 273
    Implications 274
        Driving or Supporting? 274
        Necessity of Redesign of Diploma Exam 275
        Method of Score Reporting Requires Revision 275
        More Validity-based Research Required 276
        Critical Assessment Literacy Required 276
    Research Implications 277
        National Comparative Study Needed 277

Research Across all Grade Levels is Necessary 278

Renewed Research into Writing Assessment Design 278

Limitations and Promises 279

Bibliography 281

Appendix 1: Grade 12 Writing Requirements 298

Appendix 2: English 30-1 Diploma Exam (June 2005) 301

Appendix 3: Student Survey 316

LIST OF TABLES

Table 1.1. Guba and Lincoln's Critiques of Positivist Inquiry and their Implications for Qualitative Inquiry 23

Table 1.2. Denzin and Lincoln's Five Essential Differences Between Qualitative and Quantitative Research 24

Table 1.3. Lincoln's Seven Criteria which Describe Validity as and Ethical Relationship 27

Table 1.4. Maxwell's Five Categories of Validity in Qualitative Research 28

Table 2.1. Maxine Hairiston's Description of Advances in Composition Pedagogy 105

Table 2.2. Important Differences in Processes and Cognitive Orientation Between Strong and Weak Writers 106

Table 4.1. Strengths of Quantitative and Qualitative Research Methods 137

Table 4.2. Overview of Data Collection Methods 138

Table 6.1. Comparison of Sample and Population Data (in percentages) 246

Table 6.2. Intensity of Student Experiences with Various Teaching Strategies in English 30-1 247

Table 6.3. Students' Beliefs about Good Writing 248

Table 6.4. Student Interpretations of the Importance of Classroom Writing Assessment Constructs to their Performance on Classroom Assessments 250
Table 6.5. Student Interpretations of the Importance of Process Related Elements of Classroom Assessment Constructs to their Performance on Classroom Assessments 251
Table 6.6. Student Interpretations of the Importance of Extraneous Construct Elements to their Performance on Classroom Writing Assessments 252
Table 6.7. Student Interpretations of the Importance of Extraneous Construct Elements to their Performance on the English 30-1 Diploma Exam 254
Table 6.8. Student Interpretations of the Importance of English 30-1 Diploma Exam Construct Elements for their Performance on the Diploma Exam 255
Table 6.9. Student interpretations of the Importance of Process Related Elements of the Diploma Exam Construct to their Performance on the Diploma Exam 256
Table 6.10. Student Writing Habits 264

LIST OF FIGURES

Figure 3.1. Construct Representation in Alberta's English 30-1 Writing Exam 128

CHAPTER 1: FOUNDATIONAL PERSPECTIVES

How do I know what I think until I see what I say?
- E.M. Forester

Writers write to influence their readers, their preachers,
their auditors, but always, at bottom, to be more themselves.
- Aldous Huxley

I have often confessed to my students—whether they be English majors in a teacher education program or grade 12 English language arts students on the cusp of completing high school—that evaluators can never eliminate subjectivity when assessing writing: they can use tools such as rubrics to channel their subjectivities; they can be open about the qualities or perspectives in writing that they value; but they can never be purely be objective.

This same understanding is widely accepted within the field of qualitative research. Qualitative researchers acknowledge that their work is shaped by experiences, perspectives, and theories that accompany the researcher as he or she enters the research.

Just as teachers need to be explicit about subjectivity, I will begin by being explicit about several of the key experiences that have shaped my perspectives on the question that this study attempts to address: I begin by revealing a number of my personal experiences related to the teaching and assessment of writing; I then explore critical understandings of historical issues in writing assessment; and I conclude with an examination of important issues related to the ontological and epistemological perspectives which inform this research.

Writing My Life: Selected Episodes of My Personal Journey

My family, especially on my mother's side, is an artistic family. My grandparents love art; their home has always impressed itself upon me as being a cozy gallery, appointed with stately antique rosewood furniture and adorned with original artwork from Europe and North America. My grandmother paints and plays the piano; my grandfather designs beautiful gardens. Many of my mother's siblings are skilled photographers, painters, or pianists. My mother inherited her father's gift for gardening and her mother's talent with crafts and music. She creates beautiful needlework and vibrant gardens, and she is a skilled tailor and pianist.

My father, while decidedly not artistic, is an able communicator, adept at writing and public speaking. His father is a gifted storyteller. As a child, I vividly remember being regaled by my Opa with stories of his childhood, of his exploits in the Dutch resistance during the Second World War, and of his experiences as an immigrant to Canada. My Opa always kept a journal, faithfully updating it each evening for as long as I can remember. At the age of 81, he transformed those journal entries into a moving biography (written entirely in English, his second language) of the first sixty years of his life. My father, it seems, has acquired his father's penchant for communication. He works as a pastor, and as such spends much of his time public speaking and writing. He is a good writer—clear, concise, and in tune with the needs of his audience.

I suppose it is from the confluence of these two sets of influences, my mother and her heritage and my father and his, that I learned early in life to value the beauty of art and the power of personal expression.

While I learned to value both visual and verbal forms of expression, I also discovered that my abilities rested in mostly one of these two realms. I remember clearly the moment when I was first confronted with the question, "What means of self-expression best suits my abilities?" My uncle was at the kitchen table sketching an image of a knotted old man fly-fishing; my brother was a few feet away picking at his guitar as he sang a Kenny Rogers' tune. Watching, listening, I felt saddened, incomplete. I couldn't draw, paint, or play a musical instrument, yet I craved a similar means of expression. I knew that I enjoyed writing. It had been a while, however, since I had been asked to write anything creative or expressive in school, so I decided to take up the challenge on my own, to develop my capacity to express ideas through creative writing. In high school, I began writing poetry, short stories, and the occasional personal essay. Since then, much of my academic and professional career has revolved around developing my skills as a writer, my ability to assist others with their writing, and my understanding of the processes through which writers write. This focus has led me to the question that rests at the heart of my doctoral research.

Steps on the Journey

Grade Eleven: Directions

The tedium of my junior and senior high education was beginning to wane as the prospect of a future beyond high school presented itself.

Our Biology lab was new. The rows of tables were yet unmarred by students' personalized engravings. The white boards had not yet acquired the black smudge that seems the hallmark of their use. Our teacher was rested and enthusiastic about his lectures. As he talked, my classmates feverishly scrawled their notes. I too was busy scratching at my notebook. Only, I was not taking notes, I was "Awriting." I had determined, rather naively, during the previous school year, that I would become a writer. At the time, I had planned to one-day write a great Canadian novel. To start with, however, I wrote poetry, a few short stories, and a trite novella about a man who won the lottery, secretly cashed in his ticket, left his wife, and traveled the world while she hunded him down

Grade 12: Writing for the Provincial Exam

I learned more about how to write from my father than I did from my high school English teachers. My high school teachers were very much old paradigm writing teachers: They assigned topics to write on, they gave time in class to plan our assignments, and they asked for final drafts to be handed in for grading. They seldom discussed the processes of writing, they rarely ever conferenced with us, and they focused their grading mostly on issues of correctness. Papers were returned after a week or two with a few comments, a range of corrections, and a mark. The mark was always out of twenty or twenty-five; it did not come with a scoring guide or anything that provided explanation or justification. The mark itself appeared, as if out of nowhere.

It was my father who conferenced with me. He challenged me to think about audience, to go beyond simplistic ideas, and he prompted me to think more broadly.

After first drafts were completed, he helped me think about revisions. After later drafts were completed, he helped me understand syntax, grammar, and rules of punctuation. Through my work with him, I learned the value of multiple drafts, of revising for clarity, and of patience when the writing was not going as well as I would have hoped. My writing abilities grew because of him.

In British Columbia, at the end of grade 12 students were expected to write a number of exams, an English exam being one of them. Part of this exam required students to respond to a general prompt with an essay or narrative that somehow touched on the prompt's subject matter. The first idea that came to my mind after reading the prompt was tangentially related to the topic. I decided to run with it. After having completed the essay, however, I discovered that I had little time to go back over the essay to rethink my choices.

Several months later (it was the middle of summer, and I had already been preparing for my trip to Alberta to begin my first year of university), I received my marks report in the mail and was stunned to find that though I regularly received marks in the mid nineties from my English teacher, I had received a mark in the low eighties on my government exam. I puzzled over why this might be the case, but having received no additional information (other than the score on the exam), I decided not to dwell on the disappointment. I had already been accepted into university and was moving forward, not looking back.

Undergraduate Choices

The first major choice in my undergraduate program was decided for me. I had originally wanted to complete an honors BA in creative writing, but because I did not have the right pre-requisites on my high school transcript, I was not able to enroll in a BA program. Instead, I was accepted into a BEd program where I majored in English and minored in Teaching English as a Second Language. In spite of this shift in focus, many of my English courses, and virtually all of my option courses, were focused on writing.

In terms of my development as a writer, WRITE 398 *Intermediate Creative Writing: Nonfiction* was perhaps the most important course I completed as an undergraduate student. The textbook that was used in this course was Richard M. Coe's (1990) *Process, Form, and Substance: A Rhetoric for Advanced Writers.* This book helped me to understand writing as a dynamic process. Until this point, I had engaged in a rigid writing process: I researched, planned, drafted (from introduction to conclusion), revised, and polished. My father certainly helped me with each of these stages, but I was often terribly frustrated during the early stages of writing, especially when it came to drafting.

One particularly painful writing experience that occurred early in my undergraduate program illustrates well the frustrations that I often faced when writing. For an English literature course, I was asked to write an essay in which I critically analyzed the novel *Love Medicine* by Louise Erdrich (1984). I read the novel, conducted an analysis, developed an outline, and began to create a first draft. I began writing at four in the afternoon, but I could not get the introduction right. I spent hours writing one introductory paragraph after another, but it seemed that each introduction led the paper into topics or ideas that I did not want to pursue. Sometimes this became clear while the paragraph was being written; at other times it became clear two or three paragraphs later. Regardless, each time this occurred I would erase the work I had done and I would start over again. I finished my first draft well after midnight. During that whole evening, it never dawned on me to forget about the introduction and start in the middle. I had been taught that you start at the beginning and worked to the end. I had also never heard of simple techniques such as freewriting or ink shedding and so it did not occur to me to just write and worry about the structures and focus later. Coe's book introduced me to these ideas. He helped me understand that I could approach each writing task differently, and he introduced me to strategies that I could use to help deal with problems that frequently occur at various stages of writing. Because of this course, my writing improved and my frustrations with writing dramatically diminished.

During the next year, I enrolled in my Advanced Professional Term. In addition to a 10-week field placement, this term included a course that focused extensively on methods for teaching English language arts. A major assignment within this course required students to develop a writing portfolio that included a reflective essay on how I would teach writing. I titled my portfolio "Stand Back: Teaching Writing from the Sidelines." I borrowed the notion of standing back from Donald Murray (1968) who argued that the most important goal of a writing instructor is to help students develop confidence and independence in their writing, to wean them off their dependence on the teacher, and to help them become more reliant on themselves and their peers. This notion of standing back carried with it a number of implications. When assigning writing, it implied that the teacher should provide students the freedom to choose their own topics. When providing feedback to students, it implied asking questions until students began formulating answers, rather than identifying errors and asking students to correct them. Combining this idea with what I had learned from Coe, standing back meant providing students with strategies for working through complications that presented themselves at various stages of the writing process; standing back meant not predetermining the process students were to follow when writing; it meant creating in my classroom the space for students to individually engage in their own processes. It meant providing space for conversation. Standing back meant conferencing with students in a manner that enabled them to remain in control of their writing as opposed to responding to my vision of what their writing should be.

Beginning Teacher, Finding my Way

I began my career teaching in a rural K-12 school. Essentially, as a first year teacher I was the junior and senior high English department. I felt inadequate to the task. Among other struggles, I was lost over how best to teach writing. Questions I struggled with as I prepared my year plans included: What did students most need to know about writing? How should I structure and focus my writing instruction? And, what approach to teaching writing works best?

I knew, based on my undergraduate teacher education, that I wanted to teach writing within a process oriented approach, one that utilized student and teacher conferences as a primary teaching tool and one that had students complete a writing portfolio during their school year. I quickly discovered, however, that this approach was not as easily implemented as I might have hoped. For example, student portfolios seem like a much better idea when used with fourth year education students than they do when used with a group of junior high students, half of whom seem unable to remember their textbooks from day to day, let alone keep track of the assignments they have been working on during the term. My first attempt at portfolios failed because of organizational issues (my own, and my students'). I also found a conferencing approach to be more difficult than I had expected it to be. I struggled with managing a classroom environment, keeping twenty to thirty students on task, while at the same time working with individual students. The focus required to conference made it difficult for me to pay attention to how much or how little work my students were doing. I also found conferences to be inefficient: In a class of thirty, it could take me over a week to conference with each student.

I tested two other ideas about writing pedagogy that year and found each of those wanting as well. I had learned from Janet Emig's (1971) work to think of the development of a piece of writing or of an idea to be organic rather than formulaic. I learned to despise the five-paragraph essay form and the simplistic idea development that accompanied this and other forms of formula-based writing. When completing my own writing, I rarely used formulas to guide my process or to provide structure to my work. I began with an idea and followed it as it blossomed into new ideas and as it provided structures and organizational patterns. However, when I tried to teach writing this way to my junior high students I often found their essays and short stories to be poorly organized, highly repetitive, and often weakly developed. I knew that structure and organization developed across drafts but because I had moved away from a conferencing style approach to teaching, I had little opportunity to review students' work in progress. Often I took their writing in for feedback and grading—writing they thought was polished but writing that was really mostly first draft writing. Consequently, many of the papers I assessed were marked by poor idea development and significant organizational problems. In desperation, I introduced the five-paragraph essay structure to my students. This helped them with paragraph-to-paragraph structure (though many problems within paragraphs persisted), but it did little to improve the quality of ideas that students were developing. Similarly, I had decided that I would teach grammar within the context of student writing, that I would not use grammar worksheets, and that I would ensure that grammar lessons were meaningful to students. I found, however, that once I stopped conferencing with students, I had a difficult time doing this. I tried to adopt a "goof sheet" approach that I had read about during my undergraduate training. In this approach the teacher records in shorthand—in the margins of a student's paper—all the grammar, punctuation, and spelling errors that the teacher finds in a student's paper. The student is then required to look up each of the errors in a writer's handbook, make a list of errors and methods of correcting them, and then correct each error. I quickly found this approach unworkable; it was far too time consuming for me as an instructor to identify every error in a student's paper; it was far too time intensive for a student to look up and correct every error in each paper. Quickly I abandoned this approach. Instead, I taught mini-lessons on common errors that students had made on previous assignments.

Achievement test and diploma exam results at the end of that first year suggested that my students were performing well. Our test results were slightly above the provincial average. I ended that first year, however, feeling frustrated. I was teaching writing according to a model that I was unhappy with, one that did not utilize approaches that I believed were important.

During the next school year my approach to teaching writing did not change much: I remained product oriented in my assessment; I seldom conferenced with my students; and I limited the amount of grammar instruction I provided and began focusing more on the literature we were studying than I did on the writing that my students were doing. During this year, however, I began to question more seriously the amount of time and energy I was putting into my planning and marking. According to the provincial assessments, my students were doing well enough, but I was frustrated. I wanted my students to perform far above the provincial standards. I wanted to teach writing in a manner that reflected better what I had learned about teaching writing. I wanted to live more and work less.

During the summer between my second and third year of teaching I got married. Consequently, my concern increased in regards to the volume of time teaching was requiring of me. During that summer, I stumbled upon an approach to teaching that I thought could help address this concern while at the same time provide me with the opportunity to teach writing according to the principles I had developed in my undergraduate training. During that summer, I developed a project-based approach to teaching.

The key to project-based approaches is to ensure that each project captures a significantly large portion of the curriculum. That the project tapped into a range of language arts, that it crossed a number of genres, and that it tapped into a sufficiently large sample of learning outcomes described in the *Senior High ELA Program of Studies*. Over the course of the full year, all outcomes should be dealt with across the full range of projects students are asked to complete. This approach significantly changed my classroom dynamic. It meant that I had large planning loads at the beginning of a unit as I developed the project and the marking rubrics that would accompany it, but it also meant that across a four-week unit I might have three weeks of rather light planning. The project carried the instructional load. Much of my teaching became student centered: As students worked on their projects, they would come to me with questions about content knowledge or about skill development and I would be able to assist them in this work. A conferencing approach naturally fit within a project-oriented approach to teaching. Because I designed my projects with student interests and aptitudes in mind, and because I provided students with the freedom to make choices that interested them, I found my students to be far more engaged than they had previously been. As student engagement increased, classroom management issues subsided. This transition freed me up to focus more on working with students individually. Conferencing meant that I could move away from teaching writing through formula approaches because I was now able to work with students on their writing, to help them discover the organizational structures that accompanied the ideas that were emerging in their writing. It meant that I was able to teach students about specific grammatical or punctuation problems that were occurring in their own writing. In addition, I was able to give them instruction in the context of improving their own work. Marking became less time consuming as well. Rather than providing extensive comments on student papers, I provided comments during conferences. My comments during marking focused not on identifying error or on justifying marks, but rather on reflecting on how a student handled an issue or concern they had raised during one of our conferences. I also found marking to go more smoothly because often I had seen a student's work several times before I had to mark it. I was familiar with the piece, I knew how it had developed, so by the time I marked the final product I had a sense already of the text's strengths and weakness. Consequently, I could be more focused in my marking.

Teaching Writing and the Provincial Writing Assessments. *The Senior High ELA Program of Studies* in Alberta is an interesting document. It provides teachers with a theoretical framework for teaching English language arts, a list of genres students are expected to study and produce, and a list of skill-based outcomes that students are expected to master during their course of studies. While the document provides general and overarching guidelines for teachers, it is decidedly lacking in terms of concrete examples of expected student performance. The abstract nature of the *Senior High ELA Program of Studies* provides teachers with significant freedom and independence in their teaching while at the same time making it difficult—especially for beginning teachers—to know exactly how to build their year plans, or what level of skill development students are expected to demonstrate at the end of each grade level.

Alberta's provincial writing assessment program, on the other hand, is far more concrete. It provides explicit examples of writing assignments, marking guides, and samples of student writing which demonstrate the skill level that is expected of students. Because of its concrete nature in comparison to the abstract nature of the program of studies, I looked to the provincial writing assessment program for guidance when planning my writing instruction during my first three years of teaching.

While preparing my writing program during my first year of teaching, I studied the English 30[1] (grade 12 academic English) writing exams and the Grade 9 English language arts achievement tests developed by Alberta Education (the provincial Ministry of Education). I designed my teaching in part around the requirements of these assessments. For example, the grade 9 writing achievement test required that students write a business letter and a short essay or narrative. I assigned many letter writing and narrative writing assignments in grade nine. The English 30 exam, on the other hand, required students to write a long essay in response to a major work of fiction or nonfiction, and a short essay based on a poem. I designed my English 30 course with a heavy emphasis on reading, analyzing, and responding to poetry and full-length fiction and nonfiction. During this year, students wrote a number of assignments that mirrored the conditions, the expectations, and the foci of previous diploma exams.

During my second year of teaching, I was selected to work as a marker for Alberta Education's English 30 diploma exam marking team. The marking team would meet for five or six days and would mark approximately 20,000 exams. My experience on the marking team helped me to refine my understanding of what Alberta Education valued in terms of writing quality, and it helped me to understand, within Alberta's English 30 context, what features of a written text were most influential in determining an essay's overall score. I discovered that organization was critical to high performance. Students who were able to come to the point quickly and who were able to sustain that argument were likely to do best on this exam. For the remainder of that second year, I had my students focus on developing these skills using formulaic approaches to writing.

---

[1] Prior to 2003 the Grade 12 academic English course in Alberta was designated English 30; when the program of studies was revised in 2003, the course name changed to English 30-1.

During my third year, however, as I began to reinvent the method through which I had been teaching, I began focusing my writing instruction on a process-oriented approach. My students wrote many papers, and through conferencing and collaboration worked those papers from rough first drafts to polished final drafts. During this time I continued to mark diploma exams for the government and I began to realize that the skill set being measured by the diploma exam was different from the skill set my students were developing through a process-oriented teaching model. The process-oriented model was designed to teach students to think primarily about the development of their ideas, to engage critically with those original ideas, and to learn how to reshape and restructure them. It focused on the skill of revising over time, of finding the organic structures that emerge alongside the ideas being developed, and it focused on accepting uncertainty and ambiguity as being essential elements of first draft writing. The diploma exam, however, focused on student ability to generate well-organized and largely correct first drafts. Of all the skills it seemed to measure, student abilities to write under pressure and to compose quickly under strict time constraints appeared to be two of the most significant. I realized that my process-oriented approach was hurting my students' exam scores rather than helping them. Students could not effectively work through this process in the time allotted to them on the exam. The exam was measuring their ability to create a coherent first draft, not an effective polished piece of writing. During subsequent years of teaching English 30, I developed two writing programs: "writing for life," which focused on developing the skills needed to manage effectively what is often a messy and difficult writing process; and "writing for the exam," which focused on writing under-pressure, quick formation of ideas and structure, and the ability to generate largely polished first draft work. This split program seemed to work well: Over the next three years my students' average performance on this exam jumped from a 65% to a 78%.

PhD Studies: Refining the Question

When I first proposed my PhD research, I was still quite impressed with the exam=s capacity to shape and inform my teaching practice. I wanted to take my experience as a marker one step further and so I proposed a PhD study that focused on marker cognitive process. I had planned to investigate the thinking process of experienced markers as they negotiated the requirements of the scoring guide in relation to the quality of the essay. My goal in this research would be to utilize understandings of marker cognitive process to further inform teacher practice. Ultimately, I had hoped to use this research to assist students and teachers in maximizing their test scores.

I came to realize, however, as I began to immerse myself in the literature on writing assessment, that I had wrongly equated improved exam performance with improved learning. The literature I had read consistently challenged that assumption (Hamp-Lyons, 2002; Huot, 1990; Mabry, 1999; Weigle-Cushing 2000; Wiggins, 1994;

Yancey, 1999). These scholars did not take a test's validity for granted. They argued, instead, that tests do not always measure what they intend to measure. While this observation connected well with my experience with Alberta's diploma exam, until then I had not thought about the larger implications of this issue. As I began to think about these implications, a series of questions emerged: If a test does not measure what it intends to measure, how meaningful is the information it provides? What good is reliability without validity? If the test influences instruction, but the test may not be measuring what it intends to measure, what are the consequences for students and teachers? With this final question in mind, I began to recognize that validity was as much a pedagogical concern as it was a psychometric one. I decided to focus my research on this issue.

In the field of writing assessment there is a long tradition of writing instructors challenging the validity of assessments and advocating for test reform. Yancey (1999) traces this history. She reports that in the 1960's and 70's, student writing ability was often measured using multiple-choice assessments. Compositionists challenged the validity of this approach to testing and instead advocated for assessments based on student's actual writing. Though some success was achieved in this regard, psychometricians successfully argued for designs that maximized the assessment's reliability. Compositionists have since been arguing for renewed reform to writing assessment design. Some argue for portfolio assessment of student writing, while others argue simply for changes to existing writing conditions and scoring procedures. Much of this debate is built around the assumption that the format of the assessment and the scoring criteria influence teacher practice and student understanding of the subject matter (Wiggins, 1994). Both George Hillocks (2002), in the American context, and Sam Robinson (2000), in the Canadian context, make it clear that these calls for reform have largely gone unheeded. They conclude that large-scale assessment of student writing in secondary education contexts in North America is largely being conducted with the use of poorly designed writing exams.

In his article, "Fighting back: Assessing the Assessments," Hillocks (2003) captures the resulting sense of frustration among North American teachers of writing. Much of this frustration he attributes to the power that faulty assessments are exerting on education. He recalls listening to a conference speaker rant for over twenty minutes about the negative effects of large-scale writing assessments and concludes that such complaints on their own accomplish very little. He suggests, instead, that if writing instructors want to improve writing assessments, they must conduct research into the effects these assessments are having on "teaching and the curriculum—that is, on how teachers and administrators decide to prepare students for testing" (p. 63). He further argues that this research should serve as the basis for informing public debates on the future of testing in education. Hoffman et. al. (2003) agree. They argue that "educators know and understand the virtues and pitfalls of high-stakes tests and assessments of student learning better than policymakers, commercial publishers, and the public. Educators need to inform the public about the consequences of such accountability systems before the price paid by students and teachers becomes exorbitant" (p. 629). Both Hillocks (2003) and Hoffman et al. (2003) argue that teachers must meaningfully engage in work that will help shape public policy. The mere expression of educators' frustrations are, however, not sufficient to credibly inform this debate. Hoxby (2002) clearly illustrates this problem. In defense of public reporting of large-scale assessment results, she writes:

> Statewide standardized tests and school report cards may be unpleasant for ineffectual educators, but they should not be controversial with parents or policymakers who want to see higher achievement. (p. 2)

Without research to demonstrate the concerns of the profession, Hoxby's statement holds much power. It effectively ends debate. According to Hoxby, teachers who complain about the assessments must, by virtue of their complaints, be ineffectual educators. Within this context, the research community must assume responsibility for moving this debate forward.

Huot (1996) suggests that we have been making strides forward: We have articulated and demonstrated through research an effective approach to the teaching of writing; the new challenge is for us to carry this momentum through into the realm of testing. It is time, he says, "to build and maintain writing assessment theories and practices which are consonant with our teaching and research" (p. 564). I envision this work involving two prongs. The first prong focuses on exploring the impact current designs exert on teaching and learning, while the second focuses on the development of new designs. If new designs are more expensive to develop and administer, it is important for researchers to demonstrate that current designs are indeed not functioning as effectively as they should be.

The goal, then, of my research is to investigate student and teacher experiences of large-scale, high-stakes writing assessment, and to use this research to inform public debate regarding the use of high-stakes writing assessment in Alberta and, perhaps, beyond. Specifically, the three questions I ask and respond to in this dissertation are as follows:

1. How does the writing component of the English 30-1 diploma exam influence students' beliefs about, understandings of, and approaches to creating effective writing?
2. In what ways does the writing component of the English 30-1 diploma exam influence teachers' pedagogical and classroom assessment choices in their English 30-1 classes?
3. What contextual factors might enhance or mitigate this exam's influence on the teaching and learning of composition?

Epistemological and Ontological Issues

Because this study crosses a number of often competing fields of discourse—case study research/survey research, issues related to curriculum and pedagogy/issues related to assessment, and post-process perspectives on writing/process perspectives on writing—each containing opposing epistemological perspectives, questions regarding which perspective on reality underlies this study are bound to arise. I would like to deal with those questions at this point.

Each of the tensions described above can be captured within the larger tension between the paradigms within which qualitative and quantitative research orientations are embedded. According to Denzin and Lincoln (2005), a number of significant differences between these two research paradigms exist: On the one hand, they suggest that qualitative researchers "stress the socially constructed nature of reality, the intimate relationship between researcher and what is studied, and the situational constraints that shape inquiry" (p. 11). Qualitative researchers emphasize the importance of context in informing and shaping the knowledge constructed through research. On the other hand, Denzin and Lincoln point out that quantitative researchers focus on the cause and effect relationships between variables. This focus and its accompanying emphasis on objectivity compel them to strip away context from their research with the goal of conducting their work in a value-neutral environment.

The differences described here stem from differences of perspective on questions of ontology and epistemology. In terms of ontological differences, Guba and Lincoln (1994) suggest that while positivist paradigms hold to the idea of an independent reality, critical theory and constructivist paradigms are built upon the belief that reality is either socially or historically constructed. The ontological perspectives held within each paradigm suggest accompanying epistemological perspectives as well. Positivists and post-positivists hold to a dualistic, and consequently, objectivist perspective. They believe that the subject and researcher is independent of one another and that through a focus on being objective, a researcher can come to know the uncompromised slice of reality that he or she is studying. Critical theorists and constructivists, on the other hand, believe that the relationship between the researcher and the subject interact in transaction with one another. They believe that the researcher and the subject both bring to the research a set of experiences, beliefs, and perspectives that shape the knowledge that is formed within the research context. They believe that the knowledge generated through research does not exist independently of the researcher. Guba and Lincoln (1994) explore these differences between quantitative and qualitative research with reference to a series of internal and external critiques of the positivist perspective (see table 1.1).

Guba and Lincoln go on to describe important differences in relation to the purpose of inquiry, the nature of knowledge, and the characteristics upon which soundness of inquiry are judged. Positivist forms of inquiry, they suggest, are motivated by a desire to generate explanations that enable researchers to predict and control phenomena; they build theories or laws through a program of inquiry that focuses on the postulation and testing of hypotheses; and they judge the quality of research based on internal and external validity, reliability, and objectivity. Each of these positions, Guba and Lincoln suggest, are built upon a notion of a stable, external reality. Constructivist forms of inquiry, on the other hand, are motivated by a desire to both understand and reconstruct individual and social constructions of reality; they aim to build consensus around increasingly complex and sophisticated understandings of the phenomena under investigation; and they judge the quality of research on the basis of its authenticity, trustworthiness, and its ability to provide a catalyst to further action. The positions are dependent upon a relativist and socially constructed view of reality.

Table 1.1. Guba and Lincoln's Critiques of Positivist Inquiry and their Implications for Qualitative Inquiry

	Quantitative	Qualitative
Internal		
Context stripping	Focus on specific variables by eliminating other variables from study may lead to incomplete data: unexamined variables may have a strong effect on outcomes	Focus on context ensures that significant but unanticipated variables (or contextual factors) are not ignored
Meaning and purpose	Attempts to understand human behavior separate from issues of agency	Attempts to understand human behavior in relation to the purpose and meanings attached to that behavior
Grand theories and local contexts	Outsider theories being tested may not resonate within local contexts	Grounds theory in local contexts
General data and individual cases	Focus on generalizations which may not apply to local cases	Focus on individual cases when building larger theories
Discovery through inquiry	Focus on hypothesis confirmation often glosses over the creative divergent thinking required to build hypothesis	Focus on the creative process inherent in hypothesis building.
External		
Theory and facts	Focus on objectivity requires that observations are understood to be independent of hypotheses	Focus on research context recognizes that facts and theories are interdependent
Limitations of theory	Focus on theory falsification as a means of describing a tentative understanding of truth	Rejects the notion that research methods can enable researchers to capture "'real' truth"
Facts and values	Believes research can be conducted in a value free context	Believes that theories reflect values and that facts are interpreted on basis of value laden theories
The inquirer-inquired dyad	Believes that ideally the researcher can conduct his/her work in a manner that does not effect research outcomes	Believes that findings emerge in the transaction between researcher and researched

These qualities lends themselves toward certain methodological considerations. The thinking behind post-positivism positions the researcher as expert and compels him or her toward quantitative methods that emphasize experimental or quasi-experimental designs, designs which focus on hypothesis testing. The thinking behind critical theory and constructivism position the researcher and participant as co-constructors of knowledge and as such compel the researcher to adopt methods which emphasize dialogue as a means of building consensus around more a sophisticated understanding of the issues or phenomena under investigation. More concretely, Denzin and Lincoln (2005) describe five key differences between quantitative and qualitative research (see table 1.2):

Table 1.2. Denzin and Lincoln's Five Essential Differences Between Qualitative and Quantitative Research

	Quantitative	Qualitative
Uses of positivism and post-positivism	Use statistical methods to discover and verify theories	Use statistical methods to locate a group of subjects within a larger population
Acceptance of postmodern sensibilities	Evaluate research on the basis of reliability, validity, bias and subjectivity	Evaluate research on the basis of verisimilitude, emotionality, personal responsibility, an ethic of caring, political praxis, multivoiced texts, and dialogues with subjects
Capturing the individual's point of view	Focus on the individual separate from his or her context	Focus on the individual within his or her lived context
Examining the constraints of everyday life	Study context in the abstract	Directly explore context
Securing rich descriptions	Focus on generalizations	Emphasize the value of rich detail

Of the differences described by Denzin and Lincoln, perhaps the second is the most significant, both in terms of the ongoing debate surrounding the issue, and in terms of one major focus of this study. Guba and Lincoln (2005) point to the significance of this issue as well, specifically in relation to evolving understandings of validity. They argue that validity, regardless of what one calls the concept, is an essential issue in research. Validity speaks to whether or not one would be able or willing to act upon the knowledge derived through any program of inquiry (whether positivist or constructivist oriented). It speaks to the confidence one has in the research findings or conclusions. Guba and Lincoln divide this concept into two arguments:

> The first, borrowed from positivism, argues for a kind of rigor in the

> application of method, whereas the second argues for both a community consent and a form of rigor—defensible reasoning, plausible alongside some other reality that is known to author and reader—in ascribing salience to one interpretation over another for framing and bounding an interpretive study itself. (p. 205)

The first argument stems from a focus on objective reality and stresses the importance of applying methods of data collection with sufficient rigor to ensure that this reality is captured as completely and fully as possible. The second argument, on the other hand, focuses on how well inferences drawn from collected data are supported through rigorous argument.

In either argument, the question of confidence remains paramount. Much recent discussion around issues of validity focus on the question "How can we know when we can be confident in the integrity of our research?"? This is a question which Guba and Lincoln (2005) argue has no final answer. Instead, they point to three perspectives on how this question is best answered.

First, based on a constructivist perspective they argue that validity is best conceived in terms of authenticity. Authenticity, they suggest, can be known or measured according to three criteria: fairness, ontological and educative authenticity, and catalytic and tactical authenticities. In this context, fairness refers to the attempts of the researcher to ensure all voices essential to the research are represented honestly. Ontological and educative authenticity refers the research's capacity to raise both local and broader awareness and understandings of the phenomena under investigation.

Catalytic and tactical authenticities refer to the capacity for the research to prompt both participants and interested others into social and political action. The concept of validity as authenticity requires researchers to represent voices within the scope of their research with integrity, and—with a view to the researcher's role as public intellectual—to raise awareness, and to call others to action.

Second, they argue for validity as poststructural transgression. This view of validity emphasizes the multifaceted nature of knowledge and the accompanying demand for research to represent multiple perspectives, voices, and experiences of phenomena under investigation. This perspective emphasizes the complex nature of social phenomena in contrast to simplistic single pane or single viewpoint perspectives on knowledge.

Third, they argue for validity as an ethical relationship, one that is open, respectful, and reciprocal. Lincoln (1995) develops this idea in detail around seven key ideas (see table 1.3).

These criteria essentially coalesce around a deep and profound concern for an approach to research that embraces ethical practice as being perhaps the most important feature of quality in research. The heightened concern for ethics as an indicator of research quality reflects (or is necessitated by) a shift in qualitative research toward more personally engaging methods. In survey research, for example, a participant might interact with the researcher for a maximum of 40 minutes throughout the entire course of the research while in a case study context the participant might be engaged in the research every day for more than a month. Given this highly interpersonal nature of qualitative research, a focus on ethical relationships becomes increasingly important.

Table 1.3. Lincoln's Seven Criteria which Describe Validity as an Ethical Relationship

Criteria	Definition
Positionality	The text is open or transparent in regards to its stance
Communitarian	The research is designed to serve the community within which it was conducted
Voice	Multiple voices are represented within the text; the researcher's voice reflects passionate participation and commitment
Critical Subjectivity	Research involves a high level of self awareness which leads to personal and social change
Reciprocity	Relationship between researcher and participants is marked by trust, caring, and mutuality
Sacredness	Relationship between researcher and participants based on equality, mutual respect, and an appreciation of the human condition
Sharing the Prerequisites of Privilege	Through our research relationships we enable others to act upon or improve their life situation

Other typologies related to describing validity considerations within a qualitative framework have also been posited. Maxwell's (1992) typology, published in the *Harvard Educational Review*, is built upon the understanding that knowledge is contextually bounded. It is premised upon the idea that validity relates to the "relationship between an account and something outside of that account" (p. 283). He describes five facets to the validity argument in qualitative research (see Table 1.4): Descriptive validity, interpretive validity, theoretical validity, generalizability, and evaluative validity.

Highlighting the differences between qualitative and quantitative research, Maxwell (1992) argues that quantitative research confronts challenges to validity primarily through research design; qualitative research confronts these challenges primarily by exploring and rejecting alternative theories or explanations of the phenomena under investigation.

Table 1.4. Maxwell's Five Categories of Validity in Qualitative Research

Validity Types	Focus	Validity Argument
Descriptive		
Primary	Accuracy of what has been observed	Would others describe the observed phenomena similarly to the researcher?
Secondary	Accuracy of what has been inferred on the basis of what has been observed	
Interpretive	Meaning of observed objects, events, and behaviors to those engaged with them.	How well does the account respect the perspective of the actors involved in the account?
Theoretical	Applicability of theory to the understanding and interpretation of observed phenomena	Do the concepts or theories used to interpret an observation legitimately apply?
Generalizability		
Internal	Ability of an account to make sense of happenings within similar unobserved contexts.	Does the theory used to understand observed phenomena provide a rationale describing both an understanding of the situation studied and potential explanations for differences in context might lead to differences in results?
External	Ability of an account to provide understanding related to how differences in context can lead to differences in results.	
Evaluative	Applicability of an evaluative framework to the objects of study	Can the evaluative framework used in the study be justified?

While the majority Maxwell's validity categories do not map smoothly onto those proposed by Lincoln and Denzin (2005), a close examination of the two typologies reveals that a similar set of concerns are present in both. Maxwell's (1992) descriptive and interpretive validity concerns, for example, can be found in each of the three validity arguments posited by Lincoln and Denzin (2005). Essentially, both typologies are founded upon the principles of rigor—observation and data collection are conducted with a view to completeness and accuracy—and openness; representation of data is conducted with a view to respectfully representing the perspectives of those involved in the research. Both Maxwell and Lincoln and Denzin would agree that validity in qualitative research is dependent on these primary features.

Focus on Validity

An important tension contained within this study turns around questions of validity. While this study is embedded within a social constructivist perspective, it examines a standardized writing exam that has been developed within the context of a positivist/post-positivist perspective. Given the context within which the exam has been developed, my critique of the exam's validity is framed primarily within a positivist perspective. Perhaps some might object to this choice. However, in terms of fairness, I believe that it is important to examine the English 30-1 diploma exam within the theoretical context within which it was developed. I also believe that if the exam were found wanting in relation to a set of standards it had not been designed to anticipate, the resulting critique would lose a degree of force; however, if the exam is found wanting in relation to standards established within the theoretical framework in which it was developed, then this critique gains force, making it less easy for developers to dismiss.

Research as Transaction

My thinking about epistemology has been greatly influenced by the work of Louise Rosenblatt (1938, 1978, 1993, 1998) and her transactional theory of reader response. Rosenblatt's work was revolutionary at the time when it was first published. She argued against two competing perspectives on literary response. The first perspective places the writer and his or her intentions at the center of the text's meaning. The second perspective ignores the author and argues that meaning resides within the text itself regardless of what the author intends. Rosenblatt, however, argues that the meaning of the text is not fixed in either the text or the author's intentions, but rather that meaning emerges in the transaction between the reader and the text. She argues that as the reader engages with a text he or she constructs an understanding of the text based on his or personality, experiences, social context, interpersonal relationships, background knowledge, preconceived ideas, habits of response, and attendant moral, religious, and philosophical codes. It is through these multiple lenses that one develops an understanding of a text. Rosenblatt's perspective on ontology reflects a belief in reality as a social construction.

Within the field of literary studies, Rosenblatt argues against a purely positivist perspective on literature: She challenges the notion that only one valid interpretation of text exists, and that the reader's responsibility is to discover that meaning through a rigorous application of method. In developing this argument she anticipates—and preemptively challenges—a major external criticism of her theory: that she is arguing for a purely relativist perspective on literary interpretation. She suggests that this is not the case, while she promotes the idea that single texts can evoke multiple interpretations, she also argues that some interpretations are more or less valid than others. She states, "The work must carry its own message to each of us. Nevertheless, the student should be led to discover that some interpretations are more defensible than others. . . . The student must take into account elements that are present, and must not add to the literature elements that are not present" (Rosenblatt, 1938, p 135). In this case, she supports a notion of validity as argumentative rigor and she articulates a perspective on validity that is framed around constructivist notions of fairness. She further develops this idea in her later work where she explains:

> To speak of the text as a constraint rather than a norm or "system of norms" suggests a relationship rather than a fixed standard. Instead of functioning as a rigid mold, the text is seen to serve as a pattern which the reader must to some extent create even as he is guided by it. The text presents limits or controls; the personality and culture brought by the reader must to some extent create even as he is guided by it. The text presents limits or controls; the personality and culture brought by the reader constitute another type of limitation on the resultant synthesis, the lived-through work of art. The reader's attention constantly vibrates between the pole of the text and the pole of his own response to it. The transactional view of the "mode of existence" of the literary work thus liberates us from absolutist rejection of the reader, preserves the importance of the text, and permits a dynamic view of the text as an opportunity for ever new individual readings, yet readings that can be responsibly self-aware and disciplined. (Rosenblatt, 1978, p 130)

She argues that the reader must closely attend to the text as he or she constructs an understanding or interpretation of it. The text itself does not rigidly construct meaning, but in relation to the personal and contextual experiences of the reader, the text co-informs the structure upon which the interpretation is built. The validity of interpretation then hinges in part on whether or not important features of text are incorporated or ignored in the evolving transaction between reader and text. If important features of the text are ignored in the transaction, the understanding developed will be, at best, incomplete. This argument reflects the notion of validity as poststructural transgression. Given the exceedingly complex, multifaceted nature of social phenomena, readers and researchers are required to consider the text through multiple lenses with the aim of constructing as robust and complete an understanding as possible. Through this element of her argument, she further challenges the positivist perspective. She argues that our interpretations are limited when we focus only on a limited number of variables, or a limited range of perspectives that relate to the phenomena under investigation.

Rosenblatt also argues for a social turn in relation to her perspective on the validity of interpretation. She argues that the articulation of difference in perspective, in conjunction with shared understandings, serves to enhance collective understandings. She connects to the constructivist focus on consensus building as a means for developing and accumulating knowledge and she ties her thinking to qualitative understandings of validity. She argues:

> Recognizing all the differences between scientific inquiry and literary interpretation, we can still adapt the concept of warranted assertibility to literary interpretation. We must indeed forego the wish for a single "correct" or absolute meaning for each text. *If we agree on criteria for validity of interpretation*, however, we can decide on the most defensible interpretation or interpretations. Of course, this leaves open the possibility of equally valid alternative interpretations as was as of alternative criteria for validity of interpretation. (Rosenblatt, 1993, p. 382)

She argues here for an openness on the part of the researcher, a willingness to accept multiple interpretations, and a desire to engage with communities of learners/knowers, who collectively construct understandings and criteria for evaluating those understandings. Flinders and Eisner (1994) agree. Arguing for a transactionally-based method of qualitative research, they state:

> What researchers see and tell is always a transactional outcome, mediated first by conceptual frameworks and methods of observation, and second by the forms of representation through which a study is reported (Eisner, 1992; Schwandt, 1993). For this reason, we do not seek in educational criticism some form of ontological objectivity. Instead we look for reasonable claims and warranted plausibilities. Our search, to paraphrase Stephen Toulmin (1982), is for sound beliefs rather than certain truths. (p. 352).

On this basis, they argue for consensual validity as a means for judging the validity of interpretations. Consensual validity, they claim, is achieved when groups of scholars working within similar fields are able to agree that explanations regarding observed phenomena reflect the understandings they have developed in their own work or experiences.

Finally, a major component of Rosenblatt's transactional theory focuses on the inter-related nature of the reader and the text, or the researcher and the phenomena under investigation. She argues that the reader's stance plays a significant role in determining the interpretation or meaning one takes from the text. In relation to stances toward reading, she describes a continuum that stretches from the efferent to the aesthetic. The efferent stance positions the reader to take an instrumental approach to reading, to develop a transaction focused on taking information or ideas away from the text. The aesthetic stance positions the reader to become more intimately involved in the text, to develop a lived transaction with the text. The reading experience emerging from an aesthetic stance is fuller, richer, and ultimately more rewarding than the experience emerging from an efferent stance. Rosenblatt makes clear that the stance one takes toward a text is not predetermined by the text itself, but rather that it is determined by the reader's choice, by the readers' purpose; one can approach a rich literary text from either an efferent (focus on discrete features of text only) or from an aesthetic perspective (focus on the lived experience evoked from the text). The implications that one takes from a text are shaped by the stance one adopts when approaching the text. From the perspective of research, Rosenblatt's theory connects to the constructivist understanding of the intimate relationship between researcher and phenomena under investigation: Findings stemming from research are in part shaped by the stance the investigator chooses to take when approaching the text.

Implications for Research Design

The principles for qualitative research expressed by Guba, Lincoln, Denzin, Maxwell, Flinders, Eisner, and (indirectly) Rosenblatt have informed the general focus and shape of this study. I draw on Rosenblatt because I consider this study to be about a text (the English 30-1 diploma exam), about individual's transactions with it, and about the impact of those transactions on teaching and learning.

Based on this perspective, this study presents a series of transactions with the English 30-1 diploma exam:

- My transaction (a critical reading of the exam, informed by research in the field of composition studies and validity theory)
- Three English 30-1 teacher's transactions (a critical reading based on their lived experiences teaching English 30-1)
- One hundred and forty-two students' (10 interviews and 132 surveys) transactions (a critical reading based on their lived experiences preparing for and writing the exam)

Each transaction develops out of the experiences, perspectives, and stance of the reader. Each is unique, yet they overlap and ultimately form an important consensus regarding the exam, its validity, and its influence on teaching and learning. In Chapters Four, Five and Six I will develop in greater detail an explanation of how each transaction was recorded; in Chapter Five and Six how each was analyzed; and in Chapter Seven their implications for students, teachers, policy makers, and researchers. In the next chapter I will flesh out some of the theoretical and contextual details important to this study.

CHAPTER 2: ACCOUNTABILITY, VALIDITY, AND COMPOSITION PEDAGOGY

The greatest part of a writer's time is spent in reading,
in order to write; a man turns over half a library to make one book.
–Samuel Johnson

The artists who want to be writers, read the reviews;
the artists who want to write, don't.
–William Faulkner

The first draft of anything is shit.
–Ernest Hemingway

This study focuses on a narrow question, related to a specific educational context, which when considered more broadly taps into one of the most significant issues in contemporary North American education: We teach, study, and research within an era of test-based accountability, or all too often, an era of test-driven education. The first section of this chapter, then, focuses on issues of accountability and validity in general. The issues and perspectives discussed in this section could be taken up by researchers or teachers in any field of education. The second section of this chapter moves from this general discussion toward a focus on validity issues and concerns that have been debated during the past four decades within the field of writing assessment.

Within these first two sections a focus on the importance of construct validity emerges. This focus is taken up in the third section which focuses on issues and debates regarding composition theory and pedagogy. The question which links discussions of construct validity and composition pedagogy is a simple question: "What does good writing look like, both in terms of product and in terms of process?" While answers to this question are shaped by numerous contextual factors, the final section of this chapter attempts to address this question, first by discussing historical shifts in our understandings of writing process, then by considering perspectives on this question from a broad range of authors and teachers of writing, and finally by engaging in ongoing debates regarding emerging understandings of writing processes.

The Question of Ethics in Test-based Accountability in Canadian Education

The emphasis on accountability in Canadian education has grown over the past decade. Much of this growth has been fueled by importing concern over American educational quality to Canada's education system (McEwen, 1995), by annual reporting of school rankings in several provinces, by ongoing media coverage of the issue (Simner, 2000), and by government communication and program development. Primarily accountability has been enacted in Canada through the development and administration of provincial achievement tests, graduation exams, and minimum competency tests (Volante, 2006). During the past decade, the increased focus on accountability has resulted in both the expansion of testing programs to all but one province and to the intensification of assessments programs within provinces (Jaafar, 2006).

In May 2006, the Canadian Educational Research Association hosted a symposium titled *Accountability in Education in Canada: Ten Years Later*. Presenters representing a number of provincial accountability systems reflected on changes in accountability that had occurred over the past 10 years. Much had changed since the 1995 themed issue of the *Canadian Journal of Education* reported on the state of accountability in Canada: Anderson (2006) and Gautron (2006) report that the mission of education has been changing from one in which provincial governments attempted to ensure access to high quality education for people in all parts of the country, to one in which they attempt to provide high quality results to all people. McEwen (2006) reports that Alberta had enhanced its accountability framework by expanding its achievement test program and by introducing the results oriented Alberta Initiative for School Improvement (AISI), a program which "provides school jurisdictions with the flexibility to address local needs and priorities, while at the same time increasing accountability requirements" (p 3). Perry (2006) comments on an intensification of assessment and reporting procedures while expressing concern that the system in Newfoundland was being overloaded by accountability demands.

The two main purposes of accountability—to certify individual student achievement and to improve systems of education—however, remain unchanged over the past 10 years (McEwen, 1995; Jaafar, 2006). Whether or not these accountability systems could live up to their purposes, however, was questioned early on. Perry-Fagan (1995) conclude her review of accountability in Newfoundland with the words:

> A system can neither improve significantly, nor be held accountable for the quality of its performance, unless it has valid and reliable information on which to base decisions for action. Yet, in the education system, collecting and using performance information for accountability have long been neglected at the provincial level. To some extent, this might be attributed to the fact that acquiring data on performance is complex, time

> consuming, and costly, and that the system fears being held up to public scrutiny. (p 74)

Similarly, McEwen (1995) voices concern over whether or not test-based accountability would narrow curriculum or lead to teaching to test. These concerns highlight the need for validity-based research in relation to test-based accountability. Unfortunately, within the Canadian context, this research appears to be nonexistent. Yackulic and Noonan (1999) critique provincial testing programs on the basis of their limited scope—they measure only a handful of curricular and learning outcomes—and call for consequential validity research to explore the impact of these accountability mechanisms on provincial systems of education. They observe, "although these accountability programs are often described as contributing to educational improvement, it is not clear how the results of large-scale testing or reporting on system indictors results in improved system or student performance" (p 5). Ten years later, it is seems clear that these calls for validity-based research have not been heeded: Volante (2006) and Crundwell (2005) claim that Ontario's assessment program has only been examined in relation to content validity, while "other forms of validity have not been provided nor examined" (Volante 2006, p. 6); McEwen (2006), Perry (2006), Anderson (2006), and Gautron (2006) similarly do not report any significant validity research having been conducted in relation to their provincial accountability systems during the past decade. More concretely, Crundwell (2005) states:

> Large scale assessments used in Canada often lack strong reliability and validity data and are often developed without consideration of accepted

> test development standards to assure psychometric soundness (Miles & Lee, 2002). Very few provinces currently provide any documentation to the public regarding the reliability and validity of their assessments, or report having data on these issues that is only available internally. While lacking this data, such assessments are still used to make decisions about individual students and the effectiveness of schools and teachers with regard to the achievement of children. (p 2)

Given the stated purposes of using accountability to improve provincial systems of education and to carry high-stakes for students, this lapse in validity research is both significant and highly problematic. Downing (1996) places this ethical lapse within a human context. He suggests, "[b]ecause the stakes in all high-stakes testing programs have profound influences on people's lives, attention should be directed toward ensuring that all test interpretations and uses are responsible and legally defensible" (p. 5).

Why has the Lapse Occurred?

The political realities of accountability in education pose significant challenges to those who are assigned to devising and overseeing the enterprise. Earl (1995) observes that politics has a short-term focus making it difficult for policy advisors to develop long-range plans. Long-range plans rarely survive from one election to another. She further observes that constantly shifting policy platforms make it difficult for teachers to buy into any specific policy: if they wait long enough policies change; if they jump in too eagerly, policies change. Either scenario provides a disincentive for teacher involvement in new policy directions.

The political nature of accountability also provides challenges in regards to test validation (Madaus, 1992; Downing & Haladyna, T.M. , 1996; Smith & Fey, 2000; and Miles and Lee, 2002). Downing and Haladyna (1996) suggests a conflict of interest occurs when testing agencies conduct validity research on assessments they have developed. If the research demonstrates that the test's validity is weak, test developers may wish to suppress, rather than publish the evidence. Downing and Haladyna point out that even in cases where external researchers are contracted to conduct validity research, the assessment agencies own the results of the validity studies and so are free to publish or suppress their findings. They suggest that "evaluation should be a routine part of the test development cycle, ideally carried out by testing specialists who are independent of the sponsoring agencies" (p. 7).

Miles and Lee (2002) and Smith & Fey (2000) argue that in situations where test-based accountability is enacted, conflict of interest in relation to test validation exists. They observe that accountability is a political concern while validity is a scientific one, and they argue that when tests are used for accountability purposes, the political concerns trump all others. Smith & Fey (2000) point out that the short term focus of political life conflicts with the length of time required to conduct serious validity research. They further suggest that negative results of validity research place political actors in troubling situations: retracting policy leaves one appearing weak or indecisive, attributes which can carry negative consequences for political leaders. Rather than risk the consequences of negative findings regarding test validation, politicians and those who serve them might prefer validity research not be conducted in the first place. Meaghan and Casas (1995) report on numerous independently conducted studies which raise serious questions about the validity of tests used within accountability frameworks. This independent research certainly reinforces the political concern. Similarly, Smith and Fey (2000) review a sample of validity research and conclude:

> From the rational expert model of psychometrics, tests in high-stakes accountability programs have, so far, demonstrated questionable validity. From the perspective of the polis, however, these programs remain a great success. Tests and the sanctions and rewards that are attached to them convey the public image of fairness, toughness, strong leadership abilities, and the fortunes of office to the policy makers who initiated the program. (p. 342)

Given this context, it should not be surprising that the politically driven accountability systems described by Volante (2006), McEwen (2006), Perry (2006), Anderson (2006), and Gautron (2006) have not been subjected to concerted and systematic validity research.

Linn (1998) and Taleporosos (1998), however, do not suggest that the responsibility for such ethical lapses rest solely at the feet of public servants; they place responsibility with provincial legislatures, ministries of education, school districts implementing testing programs and related policies, test publishers, and the measurement profession. Though Linn does suggest that those with the authority to impose test-based accountability systems on schools bear a substantial share of the responsibility to ensure validity research is conducted, he also argues that researchers in the field of education and measurement have a primary responsibility to ensure such research is both advocated for and conducted. He writes, “While it is not the task of measurement professionals to decide the policy, it is our job to clarify the issues, accumulate evidence, and help interpret the evidence for policy makers” (Linn, 1998, p. 30). Taleporosos (1998) suggests the responsibility of all parties involved to insist on an open dialogue that can facilitate the development of a national research strategy.

The Ethics of Testing

Messick (1989) identifies test validity as essentially being an issue of ethics. Simply put, a valid test is also an ethical test.

Messick claims that at the heart of all validity studies is the question, "To what degree—if at all—on the basis of evidence and rationales, should the test scores be interpreted and used in the manner proposed" (Messick, 1989, p. 5)? In formulating this question he suggests that both the proposed test use and the interpretations of test scores be justifiable on the basis of the construct—the theoretical representation of a skill or knowledge domain—which undergirds the test. A test of writing ability, for example, should reflect theoretical understandings of the skills needed to write effectively, the process involved in writing effectively, and the criterion which characterizes the product as "effective writing." Additionally, inferences drawn from test scores should be justifiable on the basis of the construct the test is designed to capture. If a student who scores 60% on a writing test is classified by test designers to be a poor writer, that inference must be attributable to the construct which under girds the test and not to other variables. The theory of writing upon which the test is built should reflect broader understandings of writing theory so that the inferences drawn from scores derived from the test cannot be called into question. Messick (1989) writes, "Using test scores that 'work' in practice without some understanding of what they mean is like using a drug that works without knowing its properties and reactions" (p. 8). According to Messick, test scores have meaning only in-so-far as they are grounded in the construct. Moss (1995) summarizes the implications of Messick's position: "Essentially it [Messick's position] would require that validity researchers provide an *explicit conceptual* or *theoretical framework* to ground the intended inference and supporting evidence" (p. 6). By implication, this responsibility falls on the test designer as well (Moss, 1992). *The Standards for educational and psychological testing* (AERA, APA & NCME, 1999) expect that,

> [t]he construct of interest for a particular test should be embedded in a conceptual framework, no matter how imperfect that framework may be. The conceptual framework specifies the meaning of the construct, distinguishes it from other constructs, and indicates how measures of the construct should relate to other variables (p. 9, 10).

In expressing this expectation *The Standards* certainly reinforce the centrality of the construct and suggest its importance in test development. The phrase "no matter how imperfect that framework may be," however, is problematic. I certainly recognize (and *The Standards* imply) the reality that many constructs in education are very complex and that our understandings of them are continuously evolving. I also recognize, however, that test developers have a responsibility to vigorously investigate and comprehensively develop these conceptual frameworks. The phrasing of *The Standards* seems to minimize this responsibility. Moss (1995) suggests this decision was likely political. At the time, test developers suggested that this responsibility would be too great for them to bear alone. She observes that Wiley (1991) attempts to address this concern by differentiating between test validation and construct validation, the latter being more comprehensive than the former. She rightly concludes, however,

> These concerns do not obviate the need for a program of validation research grounded in an explicit conceptual framework and articulated in an integrative argument that justifies (and refutes challenges to) the proposed meaning of test score. (Moss, 1995, p. 7)

Her position supports Messick's, and it imposes an ethical burden on test developers: the necessity of developing comprehensive theoretical frameworks in which to embed their tests. It is exactly this emphasis on construct-based score meaning that ties ethics to validity. Messick (1989) writes:

> One implication of the. . . formulation is that both meanings and values, as well as both test interpretation and test use, are intertwined in the validation process. Thus, validity and values are one imperative, not two, and test validation implicates both the science and the ethics of assessment. (p. 26)

This formulation effectively shifts ethical considerations in test design from being an extra-validity consideration to being an integral element of validity.

Test validation is the process through which we come to understand what test scores actually represent. A test score without accompanying test validation research is little more than a number. We might guess at the score's meaning, but without validity research we cannot ever be certain that our guess is accurate. One might possibly argue that this guess work might be acceptable in a low-stakes testing environment but as the stakes associated with a test increase, so too to does the ethical imperative of validating the test.

The provincial testing programs described by Volante (2006), Anderson (2006), McEwen (2006), and Perry (2006) all carry high-stakes for students and teachers. Grade 12 exams in B.C., Alberta, and Newfoundland each comprise either 40% or 50% of a student's final course grade. While in Ontario, receiving a passing grade on the grade 10 literacy test is a graduation requirement. For students in these provinces, the stakes associated with their assessment programs are indeed very high. Yet, in spite of the significant consequences these assessments hold for students, no systematic approach to validating provincial exams and achievement tests has been developed or implemented in any of these jurisdictions.

The Ethics of Consequential Validity

In the most recent edition of *Educational Measurement* Kane (2006), in his chapter on Validity, acknowledges that social consequences have always been tied to the concept of validity but that its exact place within a formal theory remains a point of debate. Similarly, Cronbach (1988) observes that while the consequences of assessment cannot be overlooked by those involved in test validation some scholars may choose not to include considerations of consequences in their definition of validity. (Their arguments will be discussed in the following section.) In the context of test-based accountability, the argument for including consequences of assessment within a larger validity framework is significantly strengthened. Kane (2006) observes that:

> Test-based accountability programs (e.g., "No Child Left Behind") blur the distinction between intended consequences and social consequences by adopting as their major purpose the improvement of educational outcomes for all students (Haertel, 1999). These testing programs have moved beyond the traditional monitoring role, to the use of testing as the engine of reform and accountability in education. Since these testing programs are intended to improve (or "reform") educational institutions, it seems reasonable to evaluate them as educational programs. Program evaluations include the evaluation of intended and unintended outcomes of the program being evaluated. (p. 54, 55)

The ethical imperative for conducting validity research to ensure that test scores actually carry the meaning intended, or lead logically to the inferences drawn from them, relates to provincial accountability systems' first goal of holding students individually accountable for learning. Kane makes clear that the second goal, improving provincial systems of education, carries with it a further need for validity research. He argues that this second goal speaks to an intended consequence of an accountability program. McEwen (1995) and Anderson (2006) both suggest one intended consequence of their assessment programs is to focus teaching. McEwen (1995) claims that "[w]hat is assessed becomes what is valued, which becomes what is taught" (p. 42). And Anderson, reflecting on B.C.'s grade 12 exams writes:

> This is an important accountability function. Not only do the examinations provide quality control at the individual student level, they also control curriculum content. As one commentator noted, ". . . we don't need a program to help people implement new curriculum, all we need is an examination." (p. 5)

If the intent of provincial accountability systems is to use tests to improve educational quality, it seems logical, if not imperative, that research be conducted which explores what the impact of these accountability systems really are; are they having the positive outcome intended, or are they generating negative unintended consequences?

The troubling logic in McEwen's (1995) article further reinforces the need for research on consequences of assessment programs. When describing Alberta's diploma exam program, she observes,

> [a] student's final mark in a diploma exam course is a 50:50 blend of the diploma examination mark and the school-awarded mark. This recognizes that the exam assesses only those learning outcomes in the Program of Studies that can be effectively measured in a limited time using paper-and-pencil tests. (p. 32)

In her conclusion, however, she recognizes that exams focus teaching by explicitly stating what skills and knowledge are being valued by the government. She argues that what is valued becomes what is taught and she suggests that what is not valued (the skills and knowledge not being assessed) will receive less focus in the classroom. Resnick and Resnick (1992) agree; they argue that in terms of a test's influence on pedagogy, "you get what you assess, you do not get what you do not assess, [and that it logically follows, you must] build assessments toward which you want educators to teach" (Resnick & Resnick, 1992, p. 59). Ironically, however, McEwen (1995) passes responsibility for the narrowing of curriculum onto teachers rather than on the tests which by their limited nature are expected to narrow teachers' focus. She writes,

> The challenge for Albertans will be to use the results of provincial assessments to improve educational practices that can help students learn more effectively, and to minimize such negative effects as narrowing the curriculum and teaching to the test. (p. 42)

It is troubling that in the past ten years, Alberta Education has not investigated whether or not this narrowing of curriculum has taken place. In fact, 10 years later, McEwen raises the same concern:

> there is a fine line between effective use of accountability tools to improve the education system and escalating requirements thereby raising the stakes for school jurisdictions and schools that may result in unintended negative effects. Alberta must guard against falling prey to an accountability system that detracts from potential benefits and could lead to undesirable consequences. (McEwen, 2006, p. 23)

I agree that Alberta, and all Canadian provinces, must ensure that their accountability mechanisms do not contribute to the degradation of educational quality. However, unless a systematic approach is developed and implemented to assess the validity of these high-stakes tests and these test-based accountability frameworks, this concern will continually be present. Without research into the consequential validity of such tests, we can never be certain of what impact they are exerting on teaching and learning.

Consequential Validity Within a Renewed Validity Framework

Historically, validity has been characterized as providing "information [indicating] the degree to which the test is capable of accomplishing certain aims" (AERA & NCME, 1955, p. 15). These aims include primarily the sampling of a specified body of content. Historically, this information was obtained through a combination of three methods of inquiry: content validation; concurrent validation; and predictive validation. Construct validation, a fourth method, was utilized only in extreme conditions. *The Technical Recommendations for Achievement Tests* (1955) suggest that "construct validation should be invoked when the preceding three methods are insufficient to indicate the degree to which the test measures what it intends to measure" (p. 16). Originally, construct validation was relegated to the benches and was called upon only when other forms of validation had struck out.

Messick, however, argues that content and criterion related evidence in themselves cannot support an argument for the validity of an assessment because they are dependent on construct related evidence: A test can be shown to measure the content that it purports to measure but if this content is not tied to the construct, the test can be described as lacking validity. This issue similarly effects concurrent validity. Simply because a test generates scores similar to those generated by another assessment which measures the same construct does not in itself demonstrate that the test is measuring what it intends to measure. The scores themselves have limited meaning if they cannot be directly attributed to the construct. This issue poses a challenge to predictive validity as well. A test that predicts future performance in relation to the construct does not necessarily effectively measure what it intends to measure.

In her overview of the history of writing assessment, Yancey (1999) cites Brown (1978) who makes the observation that

> [objective] tests correlate with writing ability and predict academic success; but the number of cars or television sets or bathrooms in one's family also correlate with this writing ability, and parental education is one of the best predictors there is. (p. 490)

Clearly, the capacity of a measurement tool to predict future outcomes or correlations in itself is not a sufficient indicator of validity. The capacity of the test to predict outcomes must be linked directly to its representation of the construct if the inferences drawn from the test's score are to be considered valid.

Messick argues that validity is both a unified and a multifaceted concept. Multiple types of validity evidence—predictive, concurrent, content, criterion-related, consequential—each provides different types of information regarding the validity of a test, yet at the same time each type of information is dependent on construct validation. In this regard, he argues that construct validation provides the unifying force to this multi-faceted concept.

Messick further discusses two challenges to construct-validity-evidence: construct irrelevant variance and construct underrepresentation. Construct irrelevant variance occurs when in addition to the intended construct, the tests scores are influenced by constructs that it was not designed to capture. Construct underrepresentation, on the other hand, occurs when the test fails to capture important elements of the construct.

Each challenge is significant.

If it can be demonstrated that test scores are influenced by irrelevant variance, the inferences drawn from those scores would have limited validity. For example, if "ability to create polished text within specific time constraints" is not an element of the construct "effective writing" but examinee scores are influenced by, say, time constraints we have an instance of irrelevant variance influencing score results. In such a case our confidence that test scores are essentially a reflection of student performance in relation to the construct would be limited.

Construct underrepresentation also challenges the validity of inferences drawn from a test's scores. If for example, our understanding of effective writing includes the idea that effective writing is a product of an effective writing process, but if our assessment of student writing does not allow for process, either because insufficient time is provided or because space for rough work is not provided, we have an instance of construct underrepresentation. The test will not reflect student capability in relation to our full conception of the construct. Instead our test will reflect student capability in relation to a more limited construct: "the ability to create a polished first draft."

In formulating his unitarian view of validity, Messick (1990) argues for a renewed focus on the consequential evidence for a test's validity. He writes:

> It is ironic that little attention has been paid over the years to the consequential basis of test validity, because validity has been cogently conceptualized in the past in terms of the functional worth of the testing—that is, in terms of how well the test does the job it is employed to do (Gureton, 1951; Rulon, 1946). And to appraise how well the test does its job, one must inquire whether the potential and actual social consequences of test interpretation and use are not only supportive of the intended testing purpose, but at the same time are consistent to other social values. (p. 17)

Messick proposes two avenues of investigation: the consequences of test interpretation and the consequences of test use. When we investigate the consequences of test interpretation, we investigate the social ramifications of test scores. For example if an aptitude test used for screening purposes consistently selects Caucasian males over any other identifiable group, we may have a validity concern. If it can be demonstrated that the test is influenced by construct irrelevant variance—it is influenced by cultural and gender variables that are not related to aptitude–we can be certain that the decisions based on test scores are of limited validity. In such a case, the social consequences resulting from score interpretation are not warranted by the construct the test was intended to measure.

Similarly, the issue of test use can be investigated from a consequential validity standpoint. Hillocks (2002), in an extensive study into large-scale writing assessment in the United States, observed the following:

> Truncated thinking appears as a usual classroom process in Illinois and Texas for a variety of reasons. First, teachers imitate the state assessment prompts to prepare their students for the assessment. Second, the prompt is such that no evidence is available to the writers. Third, the criteria for judging the papers do not call for evidence (only support). Fourth, support is interpreted to include statements that reiterate or expand upon claims. Fifth, benchmark papers at the highest levels of approval incorporate little, if any, actual evidence. Sixth, students study benchmark papers as models, models that exemplify vacuous thinking. (p. 201)

He demonstrates an instance where the consequences of test use challenge the validity of the test. He directly attributes students' development of limited thinking skills to a testing program that supports truncated thinking rather than the full development of ideas. In this case too, he implicitly ties consequences of test use to construct underrepresentation. Within these testing contexts, well developed ideas are clearly not integrated into the construct being tested.

Messick was clear when suggesting that test use consequences posed a challenge to a test's validity only in-so-far as it could be demonstrated that those consequences were directly attributable to flaws in the test's representation of the construct it was attempting to measure. Consequences that could not be related to construct representation, in Messick's (1989) understanding, do not pose a challenge to a test's validity. He concludes:

> [T]his form of evidence should not be viewed in isolation as a fourth

> validity type, say, of "consequential validity." Rather, because the values served in the intended and unintended outcomes of test interpretation and use both derive from and contribute to the meaning of test scores, appraisal of social consequences of the testing is also seen to be subsumed as an aspect of construct validity. (Messick 1989, p. 8)

This formulation underscores a second ethical element to test design. Social consequences of test use and of score interpretation should be justifiable on the basis of the construct the test is designed to capture. Essentially, this issue too underscores the ethical necessity that tests be designed to effectively represent the constructs they are intended to measure.

The ethical imperatives for Ministries of Education in Canada who have chosen to enact test-based accountability systems are clear: they must ensure that validity research be conducted to determine both the relevance of their test scores and the consequences of their testing programs on their educational systems. That this research has not been done points to a serious ethical lapse within these systems.

Challenges to Messick's Validity Theory

Messick's views, however, have not been entirely accepted within the psychometric community. Those who challenge Messick seem to accept his focus on construct representation while rejecting his emphasis on investigating the consequential validity of assessments. Earl (1999), for example, claims

> Experience has shown both expected and unexpected, both positive and negative consequences of large-scale assessment. While the

> misinterpretation and misuse of test results is sometimes due to a limited understanding of statistical concepts, it is sometimes due to the "high stakes" attached to them. When they are very important to individuals and institutions, or when they are associated with rewards or sanctions, test results are very susceptible to manipulation. This is less a testing issue than a political or moral issue. Any test can be corrupted. (p. 4)

Critiques such as this one attempt to minimize the assessment designer's responsibility for the impact of assessments on systems of education.

Maguire, Hattie & Haig (1994) launch a significant assault on Messick's position. They suggest that Messick's argument for consequential validity fails for several reasons:

> 1. Over emphasis on the role of the developer of large-scale standardized tests. . . .
> 2. Emphasis on large-scale systematic and planned assessments of consequential validity. . . .
> 3. Emphasis on consequential validity favors test score use rather than test development. . . .
> 4. Stress on consequential validity [which] seems to be a reaction to the increase litigation surrounding testing practice. (p. 113-115)

Over emphasis on the developer. On this first point Maquire, Hattie & Haig argue that the consequences of test use in educational settings are the responsibility of the teaching profession, not the test developers. They further argue that the ideological framework in which achievement tests are generated is predetermined by the political institutions which mandate the tests. This framework determines how test scores are to be used; the test developer does not make this determination. Both points can be refuted on the same basis. Accountability does not necessarily reside solely with one party to the exclusion of others. Users of tests have certain responsibilities; developers of tests have certain responsibilities. While these responsibilities diverge, they also overlap. Therefore, on the basis of their particular responsibilities, each party can and should be held accountable for the consequences of their designs and/or their decisions.

Additionally, it should be noted that even though political institutions determine the ideological frameworks in which assessments are employed, test developers have an ethical responsibility to ensure that their designs are appropriate to the framework that has been mandated. If it is not possible to design an assessment that functions effectively within the ideological framework determined by the political institution, assessment designers have the obligation to clearly state their concerns. It is unacceptable, I believe, to develop unsatisfactory assessments and then shrug off responsibility for the consequences of their use simply because the ideological framework within which the test was developed was predetermined by another party.

Emphasis on large-scale assessment of consequential validity. In regards to their second point, Maguire, Hattie & Haig (1994) observe that a vast majority of assessments in education are teacher designed and classroom based, not industry designed and externally imposed. They further claim that Messick's strategies for investigating consequential validity do not apply in classroom contexts. They suggest, therefore, that rather than focus energy on large-scale systematic studies of the consequences of test use, test developers should focus on the "evidential basis for score interpretation and not [on the construction of] a list of approved and disapproved uses" (p. 113).

Their argument unnecessarily muddies the waters. As Moss (1995) observes, "Messick highlights the importance of investigating the validity of a proposed use by distinguishing use from interpretation in his analysis of validity" (p. 9). The implication is simple. If a test is designed with a proposed use (or set of uses) in mind, the test developer is ethically obligated to demonstrate that the test is, or will be, valid within the context for which it was designed. The test developer is not responsible for explaining each of the contexts for which the test should not be used (Linn, 1997). Though to be fair, if the test is later marketed with a new purpose in mind, the original ethical obligations should apply in regards to this new purpose. Sheppard (1997) argues that for each new use of a test "a fresh validity evaluation is required" (p. 8).

Maguire, Hattie & Haig's second argument also ignores the contextual differences between classroom-based and external assessments. Classroom-based assessments are designed with the context for the test established and the intended use of the test framed within that context. External assessments are designed with a much more limited sense of context. While it might make sense that classroom-based assessments need not go through a rigorous process of stipulating uses, this same does not hold true for external assessments. Precisely because of this difference, these assessments should be subject to more rigorous scrutiny. This is not to suggest that teachers who design classroom assessments should not also be concerned with test use consequences, but as Moss (1995) and as the *Principles for Fair Student Assessment Practices of Education in Canada* (1993) suggest, the higher the stakes, the more pressing the issue. For example, the consequences attributable to a reading comprehension test that contributes 5% to a student's overall grade are potentially far less significant than the consequences derived from a diploma exam based on reading comprehension which constitutes 50% of a student's final grade. Flaws in the design of this exam will certainly carry greater consequences for the student, the teacher, and potentially the educational system as a whole, than would the flaws in the design of the lightly weighted reading comprehension test. It logically follows, then, that as the potential impact of the assessment increases so too should the scrutiny of that assessment.

Emphasis favors test score use rather than test development. Maguire, Hattie & Haig's (1994) third argument is that Messick emphasizes test scores over test development. They write:

> But Messick's claim that validity is an 'integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment' (1989, p. 13) tells us that what we do with the construct in a utilitarian sense is more important than what it is. (p. 115)

They offer an alternative perspective: quality assessments should be built upon sound understandings of the constructs upon which the tests are built. They suggest that

> Our success in understanding the relevant educational phenomena depends

> in a large part on our effective development of the construct and its associated theory. If there is to be a good test, then an effective test-construct relationship is imperative. For this we must have good constructs. . . . Tests must derive from such theory, and their use must be consistent with it. (p. 115)

While it is true that Messick seems to emphasize test scores over test development, this emphasis in itself does not negate his appreciation of the importance of the construct in test development. Messick, it must be remembered, anchors each facet of validity in the construct. In terms of consequential validity, he writes:

> In general, the best protection against adverse social consequences as threats to valid score interpretation and use is to minimize in the measurement process any potential sources of test invalidity, especially construct underrepresentation and construct-irrelevant variance in the test. Thus the watchword for educational and psychological measurement is to maximize empirically grounded interpretability and minimize construct irrelevancy in the test scores. (Messick 1989, p. 11)

Arguably, once the test has been developed, marketed, and administered, it is too late to address issues of construct irrelevancy in the test scores. Construct representation must be a central concern in the design process of the test. Put another way, test scores are a product of both test development and use and as such, concern over construct representation must be significant at both of these stages of the measurement process.

Consequential validity, an overreaction to legal climate. Maguire, Hattie & Haig's (1994) fourth argument is that Messick's work is merely an over-reaction to the legal climate of the times. They suggest that consequences of test use "should be moved out of the umbrella of validity and into the arena of informed social debate and formulated into ethical guidelines such as the *Principles of Fair Student Assessment Practices for Education in Canada (1993)"* (p. 115). In making this suggestion, they imply that moving consequences of test use into our conception of validity involves a simultaneous reduction of informed social debate over the consequences of test use. Moss (1995), however, refutes this argument. Citing the 1985 *Standards,* she suggests that this movement could in fact lead to more fully informed policy debates. As our understanding of test use consequences increases, and as test developers are required to provide more comprehensive data regarding the consequences of test use, policy makers will be able to make more informed decisions regarding the use of tests.

To support their contention that informed social debate can best be achieved through the development and dissemination of ethical guidelines, Maguire, Hattie & Haig (1994) refer to *Principles of Fair Student Assessment Practices for Education in Canada (1993)*. They claim, "The *Principles* deal not only with items contained under Messick's consequential validity, but speak of follow-up and redress as well. They place ethical test use properly in the arena of professional responsibility and encourage an atmosphere of openness and questioning" (p. 115).

A close reading of The *Principles*, however, calls this assertion into question. In *Section I: Developing and Choosing Methods for Assessment,* seven principles are espoused (emphasis mine):

> 1. Assessment methods **should be** developed or chosen so that inferences drawn about the knowledge, skills, attitudes and behaviors possessed by each student are valid and not open to misinterpretation.
> 2. Assessment methods **should be** clearly related to the goals and objectives of instruction, and be compatible with the instructional approaches used. When developing or choosing assessment methods, consideration **should be** given to the consequences of the decisions to be made in light of the obtained information.
> 3. More than one assessment method **should be** used to ensure comprehensive and consistent indications of student performance.
> 4. Assessment methods **should be** suited to the backgrounds and prior experiences of students.
> 5. Content and language that would generally be viewed as sensitive, sexist, or offensive **should be** avoided.
> 6. Assessment instruments translated into a second language or transferred from another context or location **should be** accompanied by evidence that inferences based on these instruments are valid for the intended purpose. (p. 4, 5)

This "should be" construction is present in each of the document's thirty-eight principles. However, at no point do the document's authors indicate what should happen if these principles are not followed, nor do they indicate what actions toward redress one should take if one is affected by any failures to implement these principles. Rather, in the introduction to the *Principles* the authors state,

> The principles and their related guidelines should be considered neither exhaustive not mandatory; however, organizations, institutions, and individual professionals who endorse them are committing themselves to endeavor to follow their intent and spirit so as to achieve fair and equitable assessment of students. (p. 2)

Ultimately, the document does not assign responsibility; test developers are free to adhere to or reject the *Principles*.

Messick, on the other hand, clearly assigns responsibility to all test development specialists. By integrating test use consequences into an understanding of validity Messick makes it clear that test developers must take responsibility for their assessments; failure to do so would constitute a breach of professional ethics.

Also, holding test designers responsible for the social consequences of their tests does not negate the responsibility of the test user, nor does it imply that in all circumstances test designers can be held accountable for improper use of their test. By including considerations of consequences of use into the domain of validity, we develop a mechanism for holding the assessments accountable. Certainly the responsibilities lie not just with the developer but also with the user. The developer, however, can insure against liability by clearly articulating the purposes and contexts within which the tests should be used. If test users use those tests for purposes other than those specified, test users can be held accountable rather than the developers. If the developers are shown to be negligent because they have not reasonably anticipated the possible consequences that flaws in a test's design might have, the developers should be held accountable.

Let me illustrate: Suppose a car manufacturer develops a new car but does not test it properly before marketing and distributing it. In the first year that the car is on the market one hundred people die in that make of vehicle. Coroners' inquests link their deaths to flaws in the car's design. The car manufacturer can be legally held liable for negligence.

Suppose, on the other hand, that a car manufacturer develops a car, tests it thoroughly, finds the design to be reliable within specified parameters of performance, and then markets the car on that basis. Jon, a sixteen year old boy, buys the car. On a trip through the mountains, he averages 180 km/h. Zipping around a sharp corner at twice the legal speed limit, he loses control, careens off the road, and in the resulting accident, perishes. A coroner's inquest ascertains what speed the car was traveling when it left the road and rightly determined that in regards to the fatality, the fault lay with the user of the vehicle and not with the company or the designer. Messick's construction of validity ensures that the parties involved in both the design and the use of tests can be held accountable. Cronbach (1988) agrees with Messick's position. He writes:

> The bottom line is that validators have an obligation to review whether a practice has appropriate consequences for individuals and institutions, and especially to guard against adverse consequences . . . . You . . . may prefer to exclude reflection on consequences from the meanings of the word validation, but you cannot deny the obligation. (p. 10)

Without this understanding of validity, it is conceivable that though individuals "should be" held accountable for the consequences that stem from the development and use of tests, no parties in fact would be. Instead, responsibility could be held in a constant state of limbo being passed back and forth between the developer and the user of the test. In such a situation, neither party is given an incentive to ensure that the test is improved so that consequences derived from the use of the test are based upon both an adequate representation of the construct and proper use of the test. Messick's construction of validity guards against this.

Further challenges to Messick. Popham (1997) challenges Messick's definition from a slightly different perspective. He argues that, historically, validity has been concerned with the accuracy of test-based inferences, not with tests themselves. He cites the 1985 *Standards.* "Although evidence may be accumulated in many ways, validity always refers to the degree to which evidence supports the inferences that are made from the scores" (AERA, APA, NCME, 1985, p. 9). As such, he challenges Messick's contention that validity refers not only to test scores, but also to test use. Linn (1997) refutes Popham's assertion. He points out that Popham has failed to acknowledge that historically validity has been understood to refer to the "*appropriateness*, *meaningfulness*, and *usefulness* of the specific inferences made from test scores" (p. 14). Because of this failure Popham improperly limits his conception of validity.

To challenge the second half of Popham's assertion, that validity has not been concerned with the tests themselves, Linn (1997) points out that

> Cronbach (1988) has argued that 'validation of a test or test use *is*

> evaluation' (p. 4) and evaluation involves much more than a determination of 'truth' or 'accuracy' (p. 4) it also involves arguments and judgments of 'worth' (p. 4) that demand attention to consequences of test uses and interpretations. (Linn, 1997, p. 14)

Clearly, while validity is related to inferences, inferences are dependent on scores, and scores are a product of the test. Validation of inferences involves the validation of the test itself.

Mehrens (1997) too appeals to Popham's assertion that validity is concerned solely with the accuracy of score-based inferences and is not a property of the test itself. To illustrate his point, he likens a test to a thermometer and explains that if a doctor uses it to check a patient's temperature finds that the patient has a fever of 105 degrees and administers a treatment based on inferences drawn from his reading of the temperature, it is not the instrument's fault if the inference proves to be faulty and the treatment has negative consequences. He concludes, "[the] same reasoning is true in education. The accuracy of an inference about the amount of some characteristic an individual has is separable from the efficacy of any treatment (or the wisdom of any action)" (Mehrens, 1997, p. 17). The problem with this analogy and with Mehren's conclusion is that it is predicated on an assumption that the thermometer is flawless in the information it reflects; but thermometers, like tests, can be flawed, providing inaccurate information. Consider the following elaboration on Mehren's analogy:

A doctor misdiagnoses a patient based on false readings from a thermometer that (for some reason) is measuring both body temperature and ambient air temperature.

Because of this false reading and the invalid inferences drawn from it, the doctor prescribes the wrong treatment and the patient succumbs to what was a treatable illness. The consequential evidence points to a problem with either the doctor's inferences or to his method of diagnosis. Subsequent investigation determines that the instrument of measurement is flawed. The thermometer is then either repaired or thrown out. A second investigation determines that 30 patients have been diagnosed using the same thermometer, this investigation discovers a range of consequences and contextual factors which either enhance or limit the consequences stemming from the use of this flawed instrument. Had this investigation not been conducted this thermometer would continue to degrade the quality of service in the hospital in which it was being used. Sheppard (1997) too argues counter to Mehrens' position. She points out:

> In the case of coaching or teaching to the test, the threat to validity is not just that use of test results does not have the intended effect on learning; it is also that a flaw in the conceptualization of the test made it susceptible to invalid score gains that then render its use invalid. Often when we examine why the intended relationship between test and outcome did not hold up, we find that some narrowness in the content framework or limitations in item format implicitly narrow representation of the construct. In a validity investigation, we don't just express a personal preference for consequences that we like or dislike. Consequences are evaluated in terms of the intended construct meaning. (p. 8)

Sheppard demonstrates that it is a denial of professional responsibility to assume that tests flawlessly reflect information. Rather, assessment specialists have the ethical imperative to hold tests accountable. Sheppard asserts that construct-based investigations into test design is the proper vehicle through which this can be accomplished.

Both Mehren's and my illustrations point to similar issues. Collectively they acknowledge a range of responsibilities in regards to the consequences stemming from the use of assessment data. This complexity in terms of responsibility suggests a further argument for conducting consequential validity research: to ensure that each party in the development and use of assessment data takes its responsibilities seriously.

Consequential Validity: What is the Evidence?

In 1998 the journal *Educational Measurement: Issues and Practice* published a themed issue focusing on consequential validity. Articles by publishers, practitioners, test designers, and validity theorists focused on questions of responsibility, existing evidence, complexity, and research methods (Yen, 1998). The authors agreed that responsibility for conducting research into the consequences of testing was shared to varying degrees among a range of stakeholders involved (Linn, 1998; Taleporosos, 1998). They also agreed that it is difficult to study the effects of testing on systems of education—this research requires arguments of causality, linking testing systems to educational effects, an argument that is difficult to conduct in a natural environment (Moss, 1998; Green, 1998; Reckase, 1998). They further agreed that existing research into the consequences of testing on teaching and learning is underdeveloped and largely inconclusive (Moss, 1998; Linn, 1998; Green, 1998; Reckase, 1998; Taleporosos, 1998; Lane, Park & Stone, 1998). Collectively they call for the development of a systematic approach to conducting research into the consequences of testing, a program that would be national in scope.

Consequences of testing have been postulated and to varying degrees investigated. Park Lane and Stone (1998) identify an extensive (though not complete) list of potential unintended consequences of testing (direct quote, bullet format is mine):

> - Narrowing of curriculum and instruction to focus only on the specific learner outcomes assessed and the ignoring of the broader construct reflected in the specified learning outcomes;
> - The use of test preparation materials which are closely linked to the assessment without making changes to the curriculum and instruction;
> - The use of unethical test preparation materials (e.g., secure assessment items);
> - Differential performance on the assessment for subgroups of students (e.g., racial, ethnic, and gender groups);
> - And inappropriate or unfair uses of test scores, such as questionable practices in the reassignment of principals and teachers. (Park, Lane, & Stone, 1998, p. 25)

Koretz and Hamilton (2006) in their review of literature on the consequences of assessment have developed a similar list. Though they also point to two positive outcomes described in the literature: a) teachers report working harder and focusing more on achievement in the context of high-stakes assessment; b) teachers report shaping their teaching practice toward addressing goals of the accountability program (such as increased innovation in teaching).

Much of the research, however, suggests that standardized assessments lead to negative consequences for students and teachers. In addition to the issues raised by Park, Lane and Stone (1998), Stiggens (1999) observes that standardized test-based accountability generally serves to raise levels of tension and frustration in schools often without providing teachers with any tangible support in terms of finding pedagogically sound ways to raise test scores. Smith and Fey (2000) report a similar range of negative effects from standardized testing; these include imposition of reductionist curriculum in poor or minority student schools, increased focus on drilling students as a means of test preparation, and increased grade retention. In fact, much of the criticism of standardized testing revolves around its impact on teaching, its seemingly pervasive power to shape teacher practice (Meaghan & Casas, 1995).

Freedman (1995), however, suggests that this concern is misplaced for two reasons. He suggests that in Asia and Europe, teaching to the test is an accepted practice. He argues that this acceptance "reflects national consensus on curricular objectives and a sense that if a curriculum is worth learning and that if the test fairly reflects that curriculum then teaching to the test is quite appropriate" (Freedman, 1995, p. 62). While correct in some regards, his argument is built upon the assumption that the tests are designed to effectively reflect curriculum and pedagogy; all too frequently this is a fatal assumption. His second argument has more merit. He suggests that "there are many superb teachers who know exactly what and how they wish to teach their children. They get on and do so, ignoring what might be on the test. They aim for and obtain higher than average outcomes" (Freedman, 1995, p. 62). Freedman does not offer any research to support his claim, though it seems plausible enough. Firmly grounded teachers, philosophically and pedagogically speaking, may not be affected to a significant degree by standardized testing.

The question for less grounded teachers becomes, does the test support development of more effective pedagogical practices, or does it encourage the development of a poorer form of practice? This question reminds us of Messick's position on validity: that construct validity forms the core of all other validity arguments. A test built soundly upon a well-defined construct, one which effectively captures the skills and knowledge contained within a set program of studies, is less likely to encourage either teaching to the test or poor teaching practice. Shepard (2006) agrees. She reflects Resnick and Resnick's (1992) argument that because we know assessments shape teaching and learning, we need to design assessment instruments which promote effective teaching and which encourage meaningful learning. Shepard (2006) argues for assessment designs that seamlessly integrate with instruction, and that reflect a similar philosophical stance to that of the teacher and the curriculum. Assessment tasks, she argues, must embody the "full range and depth of what we say we want students to understand and be able to do" (p. 639). She concludes:

> Ideally an external assessment that was well aligned with conceptually rich learning goals would have positive impacts on instruction by exemplifying significant learning targets, provide useful feedback to teachers about curricular strengths and weaknesses, and verify individual

> student's attainments. The authors of *Knowing What Students Know* (Pellegrino et al., 2001) envisioned for the future a more balanced and coherent assessment system, where formative classroom assessment would receive attention equal to that of external, high-stakes tests and where classroom and external assessments would be coherently linked to the same underlying model of learning. (Shepard, 2006, p. 639)

Elsewhere Shepard (2000) argues for a model of assessment that is founded upon a social constructivist perspective on education. She claims that the predominate method of assessment in use today was founded upon a well accepted theoretical framework, one built upon positivist theories of knowledge, and behaviorist principles of learning, which were widely accepted until the 1980s. Since then, theory underlying curriculum development and pedagogical practice have shifted toward cognitive and social constructivist theories of knowledge and learning. Assessment practices, however, have lagged behind and remain entrenched in positivist perspectives. She argues that classroom assessment (I would argue that this applies to external assessments as well) must change in two ways:

> First, its form and content must be changed to better represent important thinking and problem solving skills in each of the disciplines. Second the way that assessment is used in classrooms [or within society in general] and how it is regarded by teachers and students must change. (Shepard, 2000, p. 7)

Sheppard's argument, while important, is not significantly new. Experts in the field of writing assessment have, to varying degrees of success, been waging this argument since the 1970s.

Historical Issues in Writing Assessment

During the 3000-year history of writing assessment, a number of issues continuously surface. Hamp-Lyons (2002) identifies these as concerns regarding scorer impartiality, exam security, and the impact of assessment on what is being taught. Yancey (1999) extends this list of issues in her survey of the modern history of writing assessment. She identifies the struggle for control of assessment design as being the main issue that has shaped, and which will continue to shape, the development of writing assessments. Assessment specialists and those who employ them have largely determined the design of writing assessment while teachers and researchers in the field of writing pedagogy have consistently lobbied to have these exam designs changed. Yancey (1999) notes a number of manifestations of this struggle in the fifty-year modern history of writing assessment. First, she observes that the format of the assessments themselves have changed, shifting from objective selected-response tests to holistically scored essays produced by students, to portfolio-based assessments of student writing. These shifts were largely driven by teachers of writing who noted a discrepancy between what they were teaching and what the large-scale, high-stakes assessments were measuring. As the profession grew in terms of its understanding of writing pedagogy, it increasingly advocated for a change in the design of the assessments. Huot (1990) notes, however, that initially this movement was successful; it was only successful to the degree that reforms could be demonstrated to fit within existing standards of testing. Assessment designers insisted that new writing assessments should be highly reliable. In order to meet these existing standards of reliability, designs were developed that incorporated the following three features: they used (1) writing prompts to focus the writers, (2) anchor papers to guide the markers, and (3) methods for calculating acceptable agreement between raters (this largely was done through the use of scoring guides) (Hamp-Lyons, 2002). Huot (1990) suggests that this reform movement would not have been successful if it had not first met the reliability requirements imposed by assessment design specialists.

Over time, these new designs, too, were challenged. Wiggins (1994) contends that when acceding to the demands of the assessment specialists, practitioners sacrificed too much. He argues that the standardized, single-session, timed, impromptu format of the exams teaches students that the most important skill in writing is to get ideas down flawlessly the first time. Since the 1960s, however, writing instruction has been built largely around writing process. Mabry (1999) also critiques this exam design. She suggests that the standardized scoring guides used in writing assessment limit student choice and constrain both creativity and scope of student writing. The voices of Wiggins (1994) and Mabry (1999), among others, made explicit a demand for a more valid form of testing writing.

Portfolio assessments emerged as a response to this demand. This form of assessment was designed to function as a sorting mechanism, a check on practice, and a means of faculty development (Yancey, 1999). It differed from previous forms of writing assessment largely in its degree of standardization. Portfolio assessments allowed students to determine the topics and the types of writing they would display, they acknowledged that the role of the reader in interpreting text was far more subjective than objective, and they allowed for negotiation of scores.

A second manifestation of the conflict between assessment specialists and compositionists in the field of writing assessment rests in the emphasis on reliability over validity as a key component of test design. As noted earlier, the predominant, early focus of writing assessment design was on the reliability of the assessments. During the era of selected response tests, the predominant question shaping assessment design was, "Which measurement can do the best and fairest job of prediction with the least amount of work and the lowest cost"(Yancey 1999)? The concerns expressed in this question are primarily psychometric ones: reliability, predictive validity, and cost.

In part, the transition from objective tests to holistically scored essays was made possible through a simultaneous argument for placing the concern for validity before reliability. Further shifts toward portfolio assessment were supported largely by proposals either to maximize validity at the expense of reliability or to redefine reliability to better suit the demands of portfolio assessment (Moss, 1994). A second argument was also important in allowing for this shift to occur. This argument too was related to the issue of validity. Messick (1989) demonstrated that the central concern for test validity should be the test's construct. Within this context, predictive validity itself was no longer considered a significant argument upon which to base the validity of a test. The emphasis rather came to be placed on how well the test represented the construct it was designed to measure. Within this context, compositionists were able to argue successfully for assessment designs which appeared to better represent the construct.

A third manifestation of this struggle for control of writing assessment design could be found in the relationship of the assessment to the classroom. Objective assessments had very little to do with the daily occurrences in the classroom; rather they were instruments that stood apart from instruction. The move toward holistically scored essays was designed to address this issue. However, Wiggins (1994) notes that these assessments, too, did not adequately reflect classroom practice. Portfolio assessments, because they are based upon products students have completed in the process of their course work, are grounded directly in the classroom.

A fourth manifestation of this struggle for control of assessment design, Yancey (1999) observes, can be found in the role of the student in the assessment. Objective assessments act upon students, compelling them to choose predetermined answers. Essay exams imposed upon students specific forms and topics. They define students as producers who create on demand. Portfolio assessments involve a significant change in the role of the student: Students create pieces that are meaningful to them and select from amongst these the pieces that best represent their work.

Each of these manifestations is grounded in the deeper issue of competing philosophies of education. Robinson (2000) observes that over time curriculum theorists have moved beyond a transmission view of curriculum to a transactional view. He notes:

> [T]he underlying rationale for curriculum has changed from a transmission perspective, a concern for facts, to a transactional perspective, a concern

> for student interaction with knowledge, and occasionally to a transformational perspective, a concern for acting upon and changing society. (p. 257)

He further observes, however, that large-scale assessment practices have not adapted to better reflect new understandings of curriculum. Rather, they remain locked in behaviorist based understandings. Improving assessment designs to better reflect current theory, he notes,

> Implies fundamental shifts in thinking. At a basic level, it means using a constructivist, rather than behaviorist, model of learning. It means becoming more student-focused in teaching practices. It means encouraging self-assessment. It means sharing power with students. It means making classrooms more democratic. (p. 275)

In Canadian English language arts education, he concludes, such shifts in assessment have not yet been made.

Clandinin and Connelly (2000) cite Lageman in their discussion of educational history. He describes this history in terms of competing philosophies. He writes,

> I have often argued to students, only in part to be perverse, that one cannot understand a history of education in the United States during the twentieth century unless one realizes that Edward L. Thorndike won and John Dewey lost. (p. 185)

Thorndike could be described as a father of behavioral psychology. Palmer, Brestler and

Cooper (2001) observe that Thorndike "applied principles of learning developed in the laboratory and quantitative measurement of individual differences to create educational psychology" (p. 233). Thorndike was primarily concerned in education, with the systematic study of the products of education. Dewey, on the other hand, was largely concerned with the experience of education. As Palmer et al. (2001) explain:

> Dewey severely criticized public schools for silencing or ignoring student interests and experiences, using artificial language (perhaps about some vague future) that only served to alienate students, over relying on testing to assess student learning, differentiating students according to their presumed ability to partake in mental or manual learning instead of offering both to all, and isolating subjects from one another instead of uniting them around students' lived experience with knowledge. (p. 180)

Though I would disagree with Lageman that the battle is lost, it is certainly clear in the field of writing instruction which camp historically has held the upper hand. As Robinson (2000) notes, however, shifts in curriculum theory and pedagogy are well underway. What is yet to be seen is the degree to which these shifts influence assessment design. In this vein, he concludes,

> Certainly no one can disagree with setting standards, but the problem becomes what standards to set. And here assessment becomes political. Does one set standards that conform to a traditional transmission view of curriculum and assessment, or choose the more difficult task, the road not yet often taken: to use assessment and standards to support and enhance a

> constructivist, transactional curriculum? (p. 276)

Hamp-Lyons (2002) too looks to the future of writing assessment. With an eye to conflicts of the past and present, she suggests principles for future test designs: they must be technological, humanistic, political, and ethical. They must harness the power of computing while not allowing the use of computers to dehumanize assessment. They must consider the needs of the stakeholders. They must recognize that testing is a form of social engineering and that what and how you assess determines what you get. And, they must recognize that the power of assessment to shape individuals and institutions carries with it an ethical obligation to wield that power appropriately. Similarly, Yancey (1999) concludes her historical review of writing assessment by acknowledging the ethical dimensions of assessment design. Assessment is not a passive activity, it is political, it is active, and it is deterministic. Assessment designers, therefore, have an ethical obligation to ensure that their instruments serve education rather than hinder it.

Standardized Testing and the Teaching of Writing

The question of standardized assessment and its impact on teaching has been discussed by Robinson (2000) and Hillocks (2002) who, in an echo of Shepard's (2000) comments on assessment and curriculum in general, argue that many current forms of standardized writing assessments are limiting pedagogy in the writing classroom. They suggest that while our understanding of writing pedagogy has grown significantly in the past decades, during that same time our approach to the standardized assessment of writing has not experienced similar growth. In the following section we will explore the growth in our understanding of writing process, critiques of the process movement, and the implications of both process and post-process theory for the teaching and assessment of writing.

Growth in Research on Writing Processes

While different iterations of the history of writing process movement exist, the generally held narrative suggests that the process movement emerged in response to the product-centered pedagogy which sat at the heart of the current-traditional rhetoric. Matsuda (2003) captures this narrative as follows:

> In the bad old days of current-traditional rhetoric, the story goes, students learned modes of discourse and applied them to write their five-paragraph themes on topics assigned by the teacher, which were then graded without the opportunity to receive feedback or to revise. Then, along came the advocates of process pedagogy who emphasized the importance of teaching writing not as product but as process: of helping students discover their own voice; of recognizing that students have something important to say; of allowing students to choose their own topic; of providing teacher and peer feedback; of encouraging revision; and of using student writing as the primary text of the course. At about the same time, research on the act of composing began to appear, providing empirical support for the teaching of writing as a process. The rise of process, the story continues, led the field toward a paradigm shift, revolutionizing the teaching of composition and providing a renewed

> sense of respectability for the profession. (p. 67)

Matsuda (2003) does argue that other narratives describing the transition from current-traditional paradigm to contemporary understandings of composition processes exist and should be taken into account. His above narrative, however, does capture many of the major developmental trends over the past forty years. Regardless, my purpose in the next few pages is not to recreate this narrative in a more extensive elaboration, but rather to highlight key studies within the field of process research which were influential in shaping how we understand writing processes.

The publication in 1971 of Janet Emig's doctoral research *The Composing Processes of Twelfth Graders* marked for many composition scholars the beginning of a new era in composition pedagogy. Emig's study was the first in the field to ask a question regarding the processes of writing rather than the product of writing (Perl, 1994). Emig's (1971) study uncovered interesting comparisons between school writing and student self-directed writing. School writing was directed to the teacher, focused on literature being studied in class, involved limited class-time or student energy on prewriting, planning, rereading, or revising, was guided by a rigid set of teacher ordained rules regarding how students should write, and was evaluated with a focus on the "accidents rather than the essences of discourse." On the other hand, self-sponsored writing, she observed, was mostly directed to peers, focused on a range of stimuli (most prominently self and human relations), involved little prefiguring but was more readily reread and revised. She further noted, that the students' self-sponsored writing processes aligned more readily with the processes being described by professional writers of the day, while school writing processes aligned more with an out-dated mode of thinking about writing, one which ascribed to the "exemplars in the secretary guides of the late eighteenth century" (Emig, 1971, p. 98). She suggests that a primary reason for this state of affairs lay with teacher ignorance regarding more progressive ideas about writing and with the fact that few teachers themselves engaged in any writing. She concluded, then, that the writing processes exemplified by student writers in school was limited in comparison to self-directed writing which was more natural and which better reflected the writing processes of experienced writers.

Donald Graves (1975/1994) built upon the work of Emig with an investigation into the writing processes of seven-year-old children. Based on his extensive work with children-writers in four classrooms he concluded that variability in student writing processes stemmed from a range of sources including students' developmental levels, their gender, their personalities, their rhetorical contexts, and their knowledge of coping strategies.

Once the idea was established in the literature that writing processes varied under a range of conditions or contexts, a series of studies comparing the writing processes of weak and strong (or developing and experienced) writers emerged. Sondra Perl's (1979/1994) study of unskilled college writers demonstrated that weak writers consistently utilized a complex composing process, which moved with predictability from prewriting to writing to editing. The process she describes, however, was not a strictly linear process. Students did not first plan their draft, then write their draft, then edit it; rather, as they composed they moved through these stages, planning what to write, putting words to paper, editing those words. This pattern (or variations thereof) repeated itself continually as students composed. Rather than helping them improve their writing, Perl (1979/1994) concludes that this process contributed to their weakness as writers because their premature and constant attention to editing disrupted the "rhythms generated by thinking and writing" (p. 55) forcing students to return to the prewriting stage in an attempt to reconstruct their line of thought. Perl also observed that these weak writers focused in their editing mostly on word or sentence level edits and did not concern themselves much with the higher order editing which effects idea reconceptualization and structural reformulation. Instead, she observes,

> Editing is primarily an exercise in error-hunting. The students are prematurely concerned with the "look" of their writing; thus as soon as a few words are written on the paper, detection and correction of errors replaces writing and revising. Even when they begin with a tentative, flexible frame of mind, they soon become locked into whatever is on the page. What they seem to lack as much as any rule is a conception of editing that includes flexibility, suspended judgment, the weighting of possibilities, and the reworking of ideas. (p. 57)

Like Emig, Perl suggests that poor pedagogy rests at the root of this problem. She suggests that teachers' emphasis on correcting surface features when grading student work causes them to envision writing as a "'cosmetic' process where concern for form supersedes development of ideas" (p. 58).

Flower and Hayes (1980/1994) investigated differences in how novice and expert writers tackled the rhetorical problems presented to them in writing assignments. Like Perl, they found that weaker or less experienced writers focused mostly on the "features and conventions of a written text" (p. 71) while more experienced writers built complex representations of their audience and their writing task, while considering their own goals, the goals of their audience, and the goals of their text itself, allowing these representations to inform their composing (and editing) processes. They further found that poor writers developed a limited sense of audience, task and purpose at the beginning of the composing process and that they did not build on it as they wrote. Stronger writers, on the other hand, continued to revise and expand their representations as they wrote, which in turn influenced their writing. They conclude that rhetorical problem solving skills are teachable, and that as students learn to develop and explore their rhetorical problems they will in turn improve as writers.

Sommers (1980/1994) too compared the composing processes of developing and experienced writers. She also observed significant difference in processes across groups underlined by a common fundamental difference in each group's understandings of writing. Developing writers, she observed, follow a composing process that is modeled after the speech process, a linear process which does not allow for revision, which considers the word as the primary unit of discourse, which considers writing to be the process through which we translate speech to prose, and which envisioned the expressed thought as a completed entity. Weak writers begin by constructing an idea of what they wish to express, then compose the text, and then rarely return to it to redevelop it, especially if the writing develops smoothly, as if by inspiration. If they do revise, they focus mostly on lexical changes. These students also focus extensively on developing their thesis statements. In an echo of Perl's comments on rigid writing process, Sommers observes,

> Since they write their introductions and their thesis statements even before they have really discovered what they want to say, their early close attention to the thesis statement, and more generally the linear model, function to restrict and circumscribe not only the development of their ideas, but also their ability to change the direction of these ideas. (p. 80)

Their rigid process prevents them from developing ideas fully or effectively, and it certainly prevents them from exploring new ideas and expanding their understandings. In a further echo of Flower and Hayes's (1980/1994) work, Sommers also observes that these weak writers do not posses the skills needed to deal with their writing as a whole. They can tinker with diction, but they lack the strategies needed to deal with their papers holistically.

On the other hand, Sommers also noted that experienced, capable writers viewed the writing process as a messy endeavor which leads to discovery, guided not by time constraints but by a sense of becoming that involves a continual process of revision. These writers, she says, embrace dissonance, searching through the confusion of their thinking for newer and better ways of understanding their topic. Consequently, these writers engage in revision in a more complex manner than student writers—they use revision as a means to clarify and create meaning. Rather than focusing primarily on lexical changes, they focus on revision at the sentence level while also attending to revisions at the higher level (paragraphs and essay structure as a whole). They understand writing to be a multi-draft process, one in which the first draft is used to discover what one wants to say, while the second draft is focused on developing a structure and an idea used to shape the piece as a whole. These writers see process as recursive, with their attentions focusing on differing issues at differing times, moving back and forth from ideas to structures.

Sommers concludes her comparison of these differing processes of student and experienced writers with a criticism of the writing pedagogy of her day. She argues that students have not learned how to revise, they have not learned to embrace ambiguity in their first draft writing, and they have not learned that revision is about re-seeing their ideas. In fact, she suggests that the writing instruction they received not only does not help them develop the tools needed to revise more effectively, but points students in the other direction:

> Current dicta on revision blind our students to what is actually involved in revision. In fact, they blind them to what constitutes good writing altogether. (p. 84)

Building on an important theme in process oriented composition research, Sommers calls for improvements in writing pedagogy.

Rose (1980/1994) investigated writing process comparing students who frequently experienced writers block to those who seldom or never did. He too noticed patterns of thinking or of composing behavior that influenced writers' degrees of success. He discovered that both "blockers" and "non-blockers" used rules to guide their writing processes. Non-blockers, however, were more flexible about these rules than blockers were; if the rule was hindering their work they ignored it or modified it, while blockers considered theses rules as absolutes. Blockers, he observed, also tended to translate heuristics into rigid rules while lacking the ability, often, to distinguish between helpful and problematic heuristics; therefore, guidelines such as "always make three or more points in an essay" become rigid rules that must be followed. Similarly, he noted that both blockers and non-blockers developed plans to help them in their writing, but that blockers' plans resembled "static cognitive blueprints" which predetermined how the text was to be built (on all levels) while non-blockers developed plans that were not plans at all, but which were open, flexible and themselves subject to revision. He suggested that blockers' problems emerge as a result of a cognitive disposition toward certainty. Blockers tend not to tolerate ambiguity while non-blockers do not feel threatened by it. Therefore, blockers develop rigid plans and rules to impose certainty and structure on a process that is often messy and confusing. He suggests that teachers of writing need to work more closely with their students, assisting them in developing a more flexible approach to writing, one which is built upon a greater appreciation for uncertainty.

The body of research described above was part of a first wave of process oriented research. This first wave was framed around an experimental, objectivist perspective. Writers were observed writing in standardized environments and under standardized conditions. They were essentially stripped of context and told to write. However, beginning in the 1980s composition researchers began to question the usefulness of this approach; they began to recognize that writers' contexts were essential to their writing.

Berkenkotter's (1983/1994) study, which compared Donald Murray's writing habits under natural versus experimental conditions, demonstrated the limitations of experimental designs in measuring writing process. Murray utilized a lengthy, recursive process during his natural writing, moving back and forth between notebook—in which he jotted ideas—drafting (on tape), editing, and reflection. In the experimental setting, however—he was asked to write an article in an empty room, on a topic he had not chosen and with which he was unfamiliar, within a one hour time frame—he was unable to produce more than two lines of text. Berkenkotter concluded that during the experimental writing task Murray's writing process was constrained by situational variables which limited his ability to complete the assigned task successfully.

Since Berkenkotter, more elaborate explorations of writing within a range of social contexts have been conducted. Russell (1999), for example, points to Rymer's (1988) research into the composing processes of eminent scientists as they write journal articles, and Kleimann's research into the processes used by Government employees, to make the point that writing within organizational settings involve "highly collaborative and highly hierarchical document cycling, with complex effects on their writing processes" (p. 83). Russell further points to the research of Devitt, Van Bazerman, Van Norstrand, and Pare to suggest that cultural systems shape both the genres and the processes of writing. He further argues that such genres and processes evolve as cultural systems and organizations change.

Writers on Writing

As the importance of context became recognized, so too did the contributions of professional writers to our understandings of writing processes. In 1971 Emig criticized the reflections of professional writers as being idiosyncratic and focused on their feelings about writing rather than their actual processes and she concluded that they were not a suitable source of data for studies of writing process. As understandings of writing process changed, so did the field's perspective on this question. Perl (1994) includes four essays by professional writers in her book *Landmark Essays on Writing Process*, arguing that the reflections of professional writers offer important insights into writing processes. In the section that follows I turn from the research on writing process to explore understandings of writing process as expressed by a range of professional writers.

The list of writers whose work I reviewed for this section includes Margaret Atwood, Joan Didion, Annie Dillard, Peter Elbow, Stephen King, Donald Murray, George Orwell, Carolyn See, Kim Stafford, and William Zinsser. The backgrounds of these writers are diverse; they are poets, essayists, playwrights, journalists, and novelists. As such, they represent a broad range of experience and perspectives on writing. Most of these authors chose to focus most heavily in their books on both the writer's motivation and the writer's processes.

Motivation

Margaret Atwood cautioned that it would be misguided to look for a single motive which drove most writers to create. She claims that "any search for a clutch of common motives would prove fruitless" (Atwood, 2002, p. xxii). George Orwell, however, in his essay "Why I Write, " suggests that writers are motivated by any combination of the following: Sheer egoism, aesthetic enthusiasm, historical impulse, and/or political purpose (Orwell, 1994, p. 92). The list is not exhaustive. Donald Murray's (1990) list of motivations extends Orwell's. He writes (1990):

> I have written for many reasons: to feed my family and myself, to get ahead, to exercise power, to call attention to myself, to be published, to understand, to entertain, to make something that is my own, to find out what I have to say, and above all, from need. (p. 4)

Orwell and Atwood do agree with Murray on one common motive for writing: the desire to understand, a desire driven by an inner compulsion. Joan Didion (1994) agrees. She concludes her discussion on writing *A Book of Common Prayer* by saying, "Let me tell you one thing about why writers write: had I known the answer to any of these questions I would never have needed to write a novel" (p. 228). Similarly, the award winning journalist, William Zinsser (1988), points out that "we write to find out what we know and what we want to say" (p. viii).

Processes

The first stage. This motivation, this search for meaning, this desire to understand, influences the writer's process. Peter Elbow (1981) divides the process into two general stages: the creative and the critical. He writes,

> I think of the open-ended process as a voyage in two stages: a sea voyage and a coming to new land. . . . The sea voyage is a process of divergence, branching, proliferation, and confusion; the coming to land is a process of convergence, pruning, centralizing, and clarifying. (p. 50-51)

While Elbow's metaphor suggests two distinct phases in writing, it is better to think about them not as a large ship moving out to sea and back, but more like a little child playing on the beach in a rubber dingy: in the span of an hour or an afternoon, he may leave and return to shore several dozens of times. So the process is not necessarily first one goes as far out to sea as possible, then one comes back, and once back the journey is ended, but rather that the movement back and forth between phases of travel constitute a larger adventure.

This first stage is a challenging one. It is dominated by uncertainty while writers struggle with understanding what it is they are trying to communicate, with what means of communication might be best, and with how best to move forward. Donald Murray (1990) describes these first moments of writing as a time of silence. Annie Dillard (1989) describes them in terms of the power of the blank page to shape and control our thinking. Murray and Dillard maintain that from these moments of silence—from the stark blankness of the page—emerge our best writing. Margaret Atwood (2002) describes the challenges:

> Obstruction, obscenity, emptiness, disorientation, twilight, blackout, often combine with a struggle or path or journey—an inability to see one's way forward, but a feeling that there was a way forward, and that the act of going forward would eventually bring these conditions for vision—these were the common elements in many descriptions of the process of writing. (p. 176)

In spite of these difficulties, these writers learn to embrace their struggles. Kim Stafford (2003) states, "this is the Tao of the writing craft: Don't fight. Don't suffer. Be with" (p.

65). He elaborates; "the feeling of not getting it is a good sign, not a paralyzing signal. The writing is hard because I am seeking connections that I did not know before—that nobody knew before" (p. 66). This search for new meaning inevitably involves risk. And it is risk, William Zinsser (1988) argues, that gives writing an edge, that elevates it from the level of cliché, and produces a vision, an idea, a perspective that is new and refreshing. But risk, Zinsser warns, involves fear. He suggests that writers learn to push their fears to the sidelines during the creative phase of writing. He claims, "we write more comfortably if we go exploring, free of the fear of not being on the right road to the right destination. . ." (p. 158). Stephen King in his book, *On Writing: A Memoir of the Craft*, echoes this claim. He writes, "I'm convinced that fear is at the root of most bad writing. . . . Good writing is often about letting go of fear and affectation" (King, 2000, p. 127-8). King argues that to minimize fear one must write for one's self, with the door closed and the world shut out. He suggests that the door only be opened once the first draft has been completed and revised.

During this initial stage the writer is the main audience for whom the writing is intended. Zinsser (1988) explains, "One of my principles is that there is no typical anybody; every reader is different. I edit for myself and I write for myself" (p. 25). Some writers, though they write for themselves, may conjure up an ideal reader (King), a hypothetical reader (Zinsser), or a primary reader (Atwood) for whom the work is intended. This initial reader is not critical but indulgent. King (2000) writes:

> Call that one person you write for Ideal Reader. He or she is going to be in your writing all the time: in the flesh once you open the door and let the

> world back in to shine on the bubble of your dream, in the spirit during the sometimes troubling and often exhilarating days of your first draft, when the door is closed. (p. 219)

The uncertainty and the struggle for meaning and understanding, coupled with the compulsion to write and the knowledge that fear and the desire to impress weakens one's writing together explains why writers must write either for themselves or for their indulgent ideal readers. To write for a critical audience from the outset of the writing process—as we ask students to do in most high-stakes writing assessments—would limit risk taking and would minimize the writer's tolerance for the unknown. Elbow (1981) emphasizes the problem of writing for a critical audience in the early stages of writing: "If you are trying to be inventive and come up with lots of interesting new ideas, it's usually the worst thing in the world if someone comes along and starts being critical" (p. 8). Carolyn See (2002) puts it:

> Write your stuff, hide it, let it stack up. Reread it. Don't worry about it. Don't look for perfection. To switch metaphors, your first writing is as delicate as a seedling. Don't show it to some yahoo who wouldn't know an orchid from a kudzu. (p. 6)

During that first delicate stage of writing, the outside world is more of a hindrance than a help. Interference at this stage shuts the writing process down.

The second stage. The second stage occurs once ideas or material are on the page. In this stage writers focus on engaging critically with the text. Kim Stafford (2003) describes the relationship between the first and second draft:

> You read the landscape of your life and find the places where grief, or fear, or sudden surprise has revealed your truth. Then by reading the best of this experience, you write down one way to tell it. Then you read your first draft to find the places where the language is most your own, most in keeping with what your life revealed. There you have it. Revision is . . . a way to carve your true voice from the words your hands have written. (p. 79)

This sense of carving or cutting is common to many of these writers' conception of this second stage of writing. As these writers search through their material, they attempt to discover what will make their work coherent. King (2000) describes his first read, after having completed his first draft and having let it sit alone for several weeks:

> During that [first] reading, the top part of my mind is concentrating on story and toolbox concerns: knocking out pronouns with unclear antecedents, adding clarifying phrases where they seem necessary, and of course, deleting all the adverbs I can bear to part with. Underneath, however, I'm asking myself the Big Questions. The biggest: Is this story coherent? And if it is, what will turn coherence into a song? What are the recurring elements? Do they entwine and make a theme? I'm asking myself, What's it all about, Stevie, in other words, and what can I do to make those underlying concerns even clearer.
>
> I want resonance. Most of all, I'm looking for what I meant, because in the second draft I'll want to add scenes and incidents that

> reinforce the meaning. I'll also want to delete stuff that goes in other directions. (p. 214)

King describes precisely the conscious awareness or critical discrimination that Elbow argues is the key to the second half of the writing process.

This search for coherence is not an end in itself; rather coherence is the door to clarity and simplicity in writing, qualities Zinsser (1988) has long claimed rest at the heart of effective writing. He argues that "only by repeated application of process—writing and rewriting and pruning and shaping—can we hammer out a clear and simple product" (p. 34-35). Similarly, Elbow (1981) advises that "in cleaning up your language you have two goals: precision and energy" (p. 134). He also suggests that a recursive process of creation and critique are required to elevate the quality of one's writing. Coherence and clarity, precision and energy are not often products of the first stage of the writing process; rather they emerge through revision and polishing. Elbow (1981) argues that to focus too early on coherence and clarity can be damaging to one's writing. He teaches that the creative process must be separated from the critical if one's writing is to flourish. He claims, "You'll discover that the two mentalities needed for these processes—an inventive fecundity and a thorough critical mindedness—flower most when they get a chance to operate separately" (Elbow, 1981, p. 7).

During this critical stage of the writing process, writers need to focus first on their ideas, what they want to say as well as the structure and mechanics of how they want to say it.

Donald Murray (1968) maintains that this second stage is where the writing occurs. During the first stage ideas are put to paper, during the second stage those ideas are shaped and formed into a meaningful piece of writing. He argues that the true craft of writing is rewriting.

While it is very helpful, I believe, to think about writing processes in terms of two general stages or phases, it is equally important to recognize that the descriptions above do not focus on a generalizable theory or set of principles, but rather as general descriptions of two important stages of writing. The openness and flexibility in regards to their processes and their approach to developing and exploring ideas reflect the observations made by the process oriented researchers discussed earlier. While this process oriented research has helped us to understand how the processes of writing can be better understood, this movement too has come under significant criticism.

The Post-process Movement

Criticism of the process movement coalesced in the 1990s around the term "post-process." While those contributing to the post-process movement come from a range of theoretical positions, Tobin (2001) suggests that their critiques of the process movement revolve around the following four themes:

1. Process pedagogy has become so regimented that is has turned into a kind of rules-driven product that it originally critiqued. (p 10)
2. Process pedagogies are irresponsible because they fail to teach basic and necessary skills and conventions. (p 11)
3. Process pedagogy is outmoded because it posits a view of "the writer" that fails to take into account differences of race, gender, and class. (p 12)

4. By focusing on the individual writer, process pedagogy fails to recognize the role and significance of context.

Each critique requires further elaboration.

Regimentation of Process Pedagogy

The regimentation of process pedagogy is generally accepted as an unfortunate reality in many school contexts. Barbara Couture (1999) for example, recalls speaking to a group of school teachers who informed her that they built their writing programs so that their students did their prewriting on Monday, drafting on Tuesday, and revision on Friday. David Russell (1999) recalls visiting his daughter's elementary classroom and seeing four one word posters—PREWRITE, WRITE, REVISE, EDIT—outlining the writing process stuck to the walls. And Tobin (1994) recounts Donald Graves' dismay when hearing teachers discuss Graves' three and four step writing processes.

Post-process theorists raise two concerns regarding this phenomenon. First, they believe that because each individual is unique, and because each attempt at composition is different from one another, there can be no, generalizeable writing process, one that is consistent across writers and across writing contexts (Kent, 1999). Second, they believe that a lock step approach to writing process misleads students into thinking about writing as a mechanical exercise.

Several explanations for this regimentation of process have emerged over the years. Tobin (2001) suggests that "regimentation has more to do with the quirks of some individual teachers and the nature of the textbook business than with some inherent flaw in the process approach" (p 11). Couter (1999) argues for a more global explanation.

She suggests that this problem is more a disciplinary problem. Historically teachers of writing (not theorists) have understood that their task has been to serve as (or up) models to their students, models which students learn to replicate as they work to improve their writing. Under the current-traditional paradigm these models center on the texts themselves. Students are encouraged to deconstruct texts and then apply the lessons they learn from these deconstructions to their own texts. When the process movement emerged, teachers refocused their modeling on the processes of others in the belief that if only students could learn to approximate the processes of expert writers they would able to write like these experts. Couture argues that teachers who teach writing this way fail to understand that a central tenant of the process movement was focused on facilitating agency, on developing ideas and voice. Consequently, their pedagogy narrowly focuses on the stages of writing process rather than on assisting students to achieve their writing goals.

Contrary to the criticism placed on them, early process researchers and theorists claim that writing is not a linear lockstep process, but rather than it is recursive and dynamic (e.g. Emig, 1971; Perl, 1979). They believe that writing is a unique mode of learning (Emig, 1971), a messy process through which a writer comes to discover, articulate and evaluate ideas.

This is not to say that deconstructing process was not an important element of the process movement. It is. And while for some, the idea may have been to develop a "big theory" (Kent, 1999), the focus has been more on helping first researchers and then students and teachers understand writing processes, so that they can deconstruct their own processes and learn from them. Kent (1999), rather than ascribing to a "big theory" believes that the process research can teach us about shortcuts which can help us move from beginning to final draft more efficiently. By implication, then, the value of process research is that it helps us to understand a broad range of shortcuts utilized by a broad range of writers. Understanding these shortcuts should not be regarded as an end in itself, but rather as a means to facilitate agency.

Failure to Teach Necessary Skills and Conventions

While it is true that the process movement shifted attention away from issues of grammar and correctness, and while it might also be true that in some cases teachers took that shift too far, it is also fair to say that the process theory itself did not advocate for such a dramatic turn. The process movement did, however, cause us to rethink our approach to teaching grammar. The decontextualized grammar worksheets of the past were frowned upon in favor of conferences and mini-lessons about grammar in the context of student writing. The pages of the *English Journal* over the past two decades demonstrate a significant focus on grammar within the process movement (e.g. Nunan, 2005; Shafer, 2004; Paraskevas, 2004; Madraso, 1993; Warner, 1993) However the range of articles also reflects a struggle on the part of practitioners to put these new theories to work. The enduring appeal of grammar worksheets is their ease of use. They are often pre-prepared and easy to mark. Teachers who are feeling unprepared to teach grammar and conventions can find grammar worksheets to be a nice crutch. On the other hand, teaching grammar in the context of student writing is a far more challenging task. Teachers who know that grammar worksheets are no longer the accepted method for teaching grammar, but who do not feel confident teaching grammar in context, may likely choose not to teach grammar at all.

What the process movement did do was shift the primary focus in composing away from issues of correctness to a focus on developing skills and strategies for putting their ideas on paper. Elbow's (1981) two stage process suggests that a focus on correctness emerges later in the process rather than earlier. But, as Coulter points out, the focus in process pedagogy is on agency, and correctness was understood to be an important element of agency: if one's writing is to achieve one's goals, that writing needs to be accepted and respected by its intended audience.

The Writer Without Context

The third and fourth criticisms of the process movement are more substantial. Early forms of research into the writing process indeed took a clinical view of the writer, probing the writer's thinking in an experimental context without explicitly considering how differences within writers or across writers' contexts might shape their writing processes. Berkenkotter's (1983) work with Donald Murray began a trend within the process movement of studying writers' processes in relation to their contexts. Kent (1999) acknowledges that process oriented scholars accept the idea of a writer as a situated individual. But he criticizes these researchers for being focused more on the processes than the context. Petraglia (1999) argues that process theory paved the foundations for post-process thinking. She observes that as a discipline, composition theory has moved beyond the simple questions of whether or not writing is a process to more sophisticated questions regarding the impact of context on processes. She points out:

> Though reductive, it nonetheless may be possible to characterize the field after process as one that is attempting to animate two complementary observations: first, that writing genres, audiences, and writers themselves are socially and culturally constructed and, second, that the ways in which writing gets produced are characterized by an almost impenetrable web of cultural practices, social interactions, power differentials, and discursive conventions governing the production of text, making writing more a phenomenon than a behavior. (p. 53-54)

The questions, then within the post-process movement focus on the individual as a socially constructed being and on the act of writing as being a contextually shaped phenomenon. While Tobin (2001) recognizes this distinction, and while he too struggled in his own teaching with putting process theory into practice he concludes:

> Still even as I criticized process theorists for not talking enough about failure, conflict, competition, resistance and the various contexts that inevitably shape texts, even as I found much of the social aspects of composing to be useful and compelling, I still found nothing that displaced or disproved the fundamental vision offered by the first process practitioners. (p. 13)

While post-process theory challenges us to ask more sophisticated questions regarding writers and their processes, it does not, as some might suggest, undermine the research on writing process developed through the 1970s and 80s, rather it challenges us to think more broadly about writers, texts, and methods of composing.

Implications for Teaching and Testing

In less than glowing terms, Janet Emig (1971) described the state of writing instruction she uncovered during her research. She suggests, "for a number of reasons, school sponsored writing experienced by older secondary school students is a limited, and limiting, experience" (p. 97). Much of this instruction, she claims, is unimodal, vague, and built upon an oversimplification of the writing process. Students were given limited time and space for composition. As a result they were unable to engage in meaningful prewriting activities, nor did they have opportunities to reconceptualize their work. Within this context, revision became a process of merely correcting minor errors.

In his contemporary study of writing instruction in the USA, Hillocks (2002) describes similar findings. He observes that most of the teachers he surveyed maintain that to some degree their teaching practice is modeled after current traditional rhetoric, an approach to writing, he argues, that leads to formulaic structures and truncated thinking. He further observes that the design of state mandated writing assessment also reflects current traditional rhetoric. Hillocks balances the responsibility for this situation. Teachers, he argues, too often are not getting the proper training needed to teach writing effectively; they are expected to teach complex skills in limited time frames, and they teach within assessment contexts that too often reinforce ineffective methods of teaching.

The body of research that emerged after Emig's study (1971) was published, while focusing on composing process, contributed significantly to emerging understandings of writing pedagogy. Its influence within the field caused Maxine

Hairston (1982) to argue optimistically that a paradigm shift was beginning to take place within the field. The former paradigm, which she labeled the "current-traditional paradigm," had been concerned primarily, she claims, with the products of writing, specifically with style, usage, and structure. It had been concerned largely with the expository essay and with modes of discourse. It envisioned the writer apart from his or her context. Its view of process was linear and rigid, envisioning writers who before they wrote knew what they wanted to say and who moved lock step from prewriting—where the primary focus was on finding an organizational structure—to writing, to rewriting—where the focus mostly was placed on lower-order editing. It also posited the idea that writing ability was somehow innate, that it could not be taught.

Based on the growing body of research, a new process oriented paradigm was developing, one which focused teachers on both the process and the products of writing, and one which understood those processes and products more thoroughly than before. Harriston (1982) ascribes twelve features to this new paradigm:

1. It focuses on the writing process; instructors intervene in students' writing during the process.
2. It teaches strategies for invention and discovery; instructors help students to generate content and discover purpose.
3. It is rhetorically based; audience, purpose, and occasion figure prominently in the assignment of writing tasks.
4. Instructors evaluate the written product by how well it fulfils the writer's intention and meets the audience's needs.

5. It views writing as a recursive rather than a linear process; prewriting, writing, and revision are activities that overlap and intertwine.
6. It is holistic, viewing writing as an activity that involves the intuitive and non-rational as well as the rational faculties.
7. It emphasizes that writing is a way of learning and developing as well as a communication skill.
8. It includes a variety of writing modes, expressive as well as expository.
9. It is informed by other disciplines, especially cognitive psychology and linguistics.
10. It views writing as a disciplined creative activity that can be analyzed and described; its practitioners believe that writing can be taught.
11. It is based on linguistic research and research into the composing process.
12. It stresses the principle that writing teachers should be people who write. (p. 124)

These understandings have helped shaped advances in pedagogical perspectives on composition. Table 2.1 captures the progression in pedagogical understandings from current-traditional approaches to process and post-process approaches.

Important to this new paradigm was the acknowledgement that writing could be taught, that negotiating writing process required certain ways of thinking and a diverse set of skills, both which could be developed. This research demonstrated that weak writers were not innately so, but that they differed from accomplished writers in significant ways (see table 2.2).

Table 2.1. Maxine Hairiston's Description of Advances in Composition Pedagogy

	Current-Traditional Paradigm	Process/post-process Oriented Paradigm
Focus:	Product Style and Usage Structure Expository essay writing Cosmetic Editing	Process – instructor support Audience Purpose Occasion Variety of modes Substantive revision
Process:	Envisions a linear process	Understands writing processes to be multiple, recursive, and contextually mediated
Writing and Learning:	Writers should know what they wish to say before they begin writing	Through writing we discover and develop new ideas and understandings
Learning to write:	Writing ability is innate	Writing ability continually evolves and develops
Studies:	Writer apart from context	Influence of context on writers and writing

The juxtaposition in the research of strong and weak writers' sensibilities and skills reinforces the understanding that writing skills can be taught, that weak writers can learn to improve if only they could learn to develop the skills and sensibilities of strong writers. With this in mind, Perl (1979) and Rose (1980) suggest that writing teachers engage in a diagnostic form of individualized instruction. They recommended that teachers work with students to decode their processes in order to determine the source of their writing problems. Once this was determined, strategies could be developed to remedy these problems. Failure to address these root problems, they argue will merely continue to lock weak writers into counterproductive processes. Flower and Hayes (1980) take this argument a step further, suggesting that struggling writers need to be taught how to diagnose their own problems. These arguments led to a new approach to writing instruction built around peer (Elbow, 1973) and teacher conferences (Calkins, 1986).

Table 2.2. Important Differences in Processes and Cognitive Orientation Between Strong and Weak Writers

Weak (developing) writers	Strong (experienced) writers
Cognitive disposition toward certainty	Embrace ambiguity in both their thinking and their writing processes
Develop rigid plans which predetermine how their writing will develop	Develop flexible plans which are subject to change when the need arises
Writing process is shaped by rigid rules; drafting focuses on surface features and conventions	Flexible writing process is discovery- oriented and thus allowed to be messy
Cosmetic editing while drafting disrupts thinking; rarely return to edit after draft completed	Recursive process focuses on different goals at separate moments: draft one discovers what one wants to say, draft two focuses on structure and idea development, etc.
Limited concept of audience and task (defined at the beginning but does not develop)	Complex representations of audience and writing task build as they write, influencing how ideas develop
Develop ideas early and do not modify, revise, or further explore them; not concerned with reconceptualizing ideas or restructuring of text	Use revision to clarify and create meaning throughout the writing process
Cannot deal with paper as a whole	Are able to determine how individual elements contribute to the paper as a whole

Researchers' descriptions of the writing processes of strong writers reflect the processes described by seasoned professional writers. They generally agree that effective writing emerges from a two stage-process. Each stage requires a different method of thinking. The first requires a purely creative orientation, the second a critical one. They agree that effective writing emerges from a process that is personally meaningful and which is prompted by a desire to come to understand more clearly. They also agree that the creative stage of writing is marked by confusion and structural chaos. This stage involves risk. As writers put their thoughts on paper they expose themselves to the criticism of the world. This risk is mitigated by writing first either for themselves or for their supportive, indulgent Ideal Reader. The second stage of the writing process is marked by a search for coherence. It involves repeated rewriting, editing, shaping and polishing.

This body of information suggests that writing teachers need to provide students with writing assignments that enable students to engage in personally meaningful writing. They need to foster a tolerance for ambiguity during their writing process, they need to encourage risk by developing a supportive collaborative writing environment within their classrooms, and they need to provide time and space within the writing program for students to engage in extensive revision.

The two dozen process oriented teachers who contributed to Coles and Volpat's (1985) book on writing pedagogy focused on the product, the motivation, and the process of writing. In terms of product, they agreed that honesty, voice, risk-taking, exploration, attention to audience, effective use of details, organic structure, and control of conventions and diction were features which marked good writing. These features, they also agreed were mutually dependent on each other. While focusing on developing student writing processes, the challenge for teachers, then, is to assist students in developing the skills needed to infuse these features into their texts. The danger suggested by the current traditional paradigm is that attention to surface features may draw students toward a focus on cosmetic editing.

The post-process research suggests that writing pedagogy needs to focus on helping the writer to understand that his or her writing and/or ideas are shaped by their cultural heritage and by the contexts within which the writing occurs. This pedagogy calls for a deep exploration of audience and purpose. For example Kessler's (2005) pedagogical approach which she calls "composing for delivery" draws on post-process perspectives. She begins by asking students to think about whom they might want to write a letter to and what they wanted to say to that person(s). She encourages students to think about writing a letter that is designed to effect change, and to think about how they might need to shape their writing so that it can achieve that goal. She requires that at the end of the assignment, students send their letters to the audience for whom it was intended. Kessler's pedagogy builds largely out of the process pedagogy but with a heightened emphasis on audience and purpose.

Bartholomea's approach to post-process pedagogy involves a departure of sorts from process-based approaches. He begins by being dismissive of students' writing (Bartholomae, 1995). He challenges students to think about their texts and their ideas, not primarily as entities they have given voice to, but rather as entities which they need to interrogate as being socially situated and constructed. He argues that in the first place their ideas are not uniquely their own, but that rather they are shaped by their social milieu. His pedagogy focuses students on engaging critically with their writing to develop an understanding of text as a socially constructed document, and to then revise and restructure it accordingly. He suggests that the role of the post-process writing teacher is to ask questions, questions which prompt students to think about their writing, questions which compel students to revise their writing. Elbow (1995) argues that Bartholomea's critique of process pedagogy is perhaps more strident than necessary. He suggests that Bartholomea's sensitivities can be accommodated within a process paradigm, for example that an expresivist view of writing can coincide with a socially constructed view of the student. Ultimately, though, Bartholomea's view of the teacher differs from Elbow's (and Murray's) in an important way—Bartholomae views a teacher who is more critically engaged with his students' work, while Elbow (at least in the early stages of writing) envisions a teacher who is more supportive, more encouraging. Tobin (2001) captures the differences in approaches with the following critique of Bartholomea's approach:

> [I]f we learned anything from Murray, Emig, and Elbow, we know that you don't teach students to write by telling them that their views on issues that concern them or their narratives about events that shaped them . . . don't count as content or count only as naïve opinions to be corrected during the course. (p. 14)

This review of literature also suggests challenges for the assessment of writing. Donald Murray's experience with the artificial writing context constructed for Berkenkotter's (1983) research provides a clear warning about the importance of the context in which writing assessments take place. Assessments must be designed to enable all students to enter the assignment from a position of confidence knowing that they can enter the topic easily, that they are free to design their response around issues or perspectives that are important to them, and to which they feel they can meaningfully talk. Murray's experience also demonstrates that even expert writers have difficulty writing within time constrained environments, suggesting that writing assessments be designed to enable each writer the opportunity engage in his or her own process (while also challenging students to develop the skills or shortcuts utilized by effective writers). The work of post-process theorists suggests that standardized writing exams should not be so standardized that they are designed with a prototypical student writer in mind. Rather such assessments should be sensitive to how students' individual contexts shape they way in which they interpret questions, the manner in which they construct ideas, and the processes through which they create texts.

In terms of scoring rubrics, the challenge for test designers is to develop scoring systems that both allow for and account for complexity and flexibility. That is, they should not be so rigidly defined that good writing is being unfairly penalized because it does not demonstrate all the features which possibly could be used to define good writing. John Mellon [in Cole and Volpat (1985)] demonstrates this necessity when commenting on a text he has chosen to represent his understanding of what constitutes effective writing:

> This essay is not artistic writing, nor does it convey a strong emotional message or a strikingly new perspective on its topic. Yet I think it is good writing, even very good writing, given its circumstances. For the essay is examination writing, test writing performed on demand to demonstrate competence in written argumentation. The college underclassman who wrote it did so in two hours on an assigned topic not announced ahead of time, with no compositional aids other than a dictionary, knowing his performance would be judged as a one-time test of his writing ability. In this test setting, the writer has produced a 600-word discussion cast in

> seven well-developed paragraphs simply begun and just as simple ended, mechanically almost flawless, and expressed in sentences of sensible content and mature form. (p. 130)

What I find striking about his argument is that it acknowledges the difference between authentic writing (to whatever degree classroom writing can be said to be authentic) and writing constrained by the exam context. In his mind, because the essay was written for an exam, it did not need to be creative, or exploratory, or infused with life—qualities virtually every other contributor to Cole and Volpat's book (1985) emphasized as being important aspects of effective writing—rather, it merely needed to be simple, organized, and mechanically flawless.

However, I believe Mellon does his students a disservice by suggesting that such lifeless writing is ever acceptable. Writing assessments need to hold students to a higher standard, one which asks them to engage in writing with a complex process, that reflects the processes of expert writers, and which challenges students to engage personally in a search for new or more complex understandings.

Validity, a Pedagogical Issue

Mellon's comments point to the main issue at the heart of this study: The design of a writing assessment is a reflection of the designer's theory of writing. This designer's theory of writing enters the classroom—an occupant within a Trojan horse—within the assessment, its silent presence within the classroom carries implications for pedagogy and learning. For this reason one can argue that validity is as much a pedagogical issue as it is an assessment issue.

This understanding will guide the remainder of this thesis: Next we will deconstruct Alberta's English 30-1 exam with a view to uncovering its construct and its underlying theory of writing. The exam's theory of writing will then be compared to the theory of writing discussed in this chapter. The final chapters of this thesis will then explore the pedagogical issues that emerge from the mismatch between these two theories.

CHAPTER 3: ALBERTA'S ENGLISH 30-1 DIPLOMA EXAM

A critic is a necessary evil, and criticism an evil necessity.
-Carolyn Wells

Writing, like life itself, is a voyage of discovery
-Henry Miller

In Chapter One I suggest that the elements of this study pertaining to an analysis of Alberta's English 30-1 diploma exam would be framed as a series of transactions between the exam and the researcher, the exam and teacher participants, and the exam and student participants. This chapter contains the first of these transactions. The perspectives which give shape to this transaction are discussed in Chapters One and Two. The first section of this chapter introduces similar analyses of standardized writing assessments conducted by other researchers. Their conclusions regarding the constructs being measured and their implications for pedagogy and learning foreshadow the conclusions developed through my transaction with Alberta's English 30-1 diploma exam.

Construct-based Challenges to Writing Assessments

Messick's view of validity provides a significant challenge to test developers in the field of writing assessment. Specifically, Messick's (1989) assertion that the ethics of testing is tied directly to construct representation is problematic for many writing assessments. It is generally accepted that the construct being measured through tests of student writing is not well understood (Hamp-Lyons, 2002), nor is it stable (Gordon et.

al., 1996).

Purves (1992) provides much of the evidence for this view. He reports on a ten year study of writing assessment conducted over fourteen systems of education. On the basis of this study he was able to draw three significant conclusions:

1. School writing is an ill-defined domain;
2. School writing is a matter of products not process; and
3. The quality of school writing is what observers report they see. (p. 109)

To support his first conclusion he notes that the construct measured by writing tests is a multi-trait construct: "writing tasks may be seen in terms of their discourse functions, their cognitive demands, and their social situations" (p. 109). Further complicating details, he notes that D'Angelo (1975), Moffett (1968), Britton, Burgess, Martin, McLeod, & Rosen (1975), Hairston (1986), and Kinneavy (1971) each have developed different schemes for subdividing the construct. He concludes, however, that "whatever the division, it is apparent that different tasks present different problems, which are treated differently by students and judged differently by raters" (Purves, 1992, p. 112).

Purves' second conclusion is essentially an extension of the first. He notes that since the 1970s writing instructors have come to understand the teaching of writing as being an issue of teaching writing process. Writing assessments, however, even within this instructional context, often focus solely on the product of writing and not the process. He suggests that this is inevitable, that we cannot assess process because we can only observe product. He writes, "How can we talk about students being better or worse planners, drafters, revisors, or editors? In each case, it seems we have to look beyond the act to the result of the act: the plan, the draft, the revision, or the edited copy" (p. 113). While the observation has some credibility, it is somewhat limited. In Alberta, for example, the *Senior High ELA Program of Studies* places a significant emphasis on student development of metacognitive skills. One expectation of this metacognitive focus is that students will be able to verbalize their processes, reflect on them, and determine methods to improve them. This idea is built upon Vygotsky's (1986) notion of inner-speech as being the foundation of intelligence. An assessment of writing that is concerned with process could tap into students' metacognitive reflections on their writing process. To support his conclusion Purves (1992) reports that when students in his study were asked to reflect on what aspects of writing were most important, an overwhelming majority most valued surface features and appearance. Those who mentioned writing process, said topic selection and editing were important. He suggests that students understood success in writing as handing in a good looking product. Based on his own experiences he supports their conclusions. He writes:

> As professional writers ourselves, we are well aware that the appearance of the manuscript is important. It is a part of the first impression that we and our peers use for the judgment of the text and the writer. . . . We in the university know that the route to promotion lies not in the writing process but in products which must be both numerous and similar to others of our imputed professions. (Purves, 1992, p. 114)

Ironically, the experiences of the professional writers I surveyed in the previous chapter flatly contradict Purves. Without exception, when they reflect upon their distinguished careers, these writers suggest that the most fundamentally important element of writing is their writing process. Purves also observes, however, "If product was important for the assessment, it was also important for the students" (p. 114). Wiggins (1994) recognizes this reality of testing as well. He questions the validity of single-sitting, timed, impromptu writing exams because they misinform students about the (un)importance of effective writing processes.

Purves (1992) does, however, note an irony in current practices of writing assessment. Designers of writing assessment seem to value polished product over effective process, yet they structure their assessments to disavow what they value. The time constraints imposed during these assessments, he writes, makes it virtually impossible for students to complete more than a first draft. He observes, "In this study as in most writing assessments and writing research studies, the product was clearly not a finished piece of prose, but a draft. One wonders whether the aim of school writing instruction is to produce quick drafters" (p. 115). Wiggins (1994) argues that this is indeed the message that we are sending to students through the current model of writing assessment.

The Exam in Context

Alberta's English 30-1 writing exam is one exam within a larger provincial assessment program. In Alberta, students write achievement test in grades 3, 6, and 9, and they write a series of diploma exams in grade 12. Alberta's Grade 12 diploma exam program is built upon three main purposes:

- to certify the level of individual student achievement in selected grade 12 courses
- to ensure that province-wide standards of achievement are maintained
- to report individual and group results (Alberta Education, 2004)

The first purpose listed above relates most directly to the construct that the diploma exams have been designed to measure. In order for these exams to serve the purpose of certification of student ability, each must effectively sample the broad range of outcomes required within their associated programs of study. In Alberta, the English 30-1 program of studies develops a comprehensive set of expectations for student writing. These expectations are listed in Appendix 1. In the paragraphs that follow I provide a brief, descriptive synopsis of these requirements.

Writing Requirements in the English 30-1 Program of Studies

Perhaps the most dominate requirement within the curriculum outcomes for grade 12 writing is the expectation that students develop the ability to assess a text in progress for a broad range of qualities including organizational components, controlling ideas, transitions, supporting details (for completeness and relevance), reasoning and logic, syntax, diction, phrasal structures, grammatical correctness and the text's ability to address audience and purpose. Students are also expected to demonstrate an ability to critically appraise and modify interpretations, perspectives and opinions. Students are further expected to reflect on experimentation with language, demonstrating how such experimentation impacts their growth as language users. They are to required to appraise their strengths and weaknesses as language users and to select and monitor strategies which they can use to increase strengths and address weaknesses. The curriculum requires that students develop and utilize a range of strategies for forming understandings and for improving language skills. It does not, however, define these strategies.

Students are expected to evaluate source material. They are expected to integrate new knowledge with old understandings, to support conclusions with relevant details and to draw conclusions relevant to findings. They are required to reflect on their writing in a broad manner; they are expected to reflect on how their sense of audience impacts their choices, how their choice of medium reflects their understanding of content and context, and on how their choice of genre compels them to address issues of content and purpose.

Given that the exam's stated purpose is to certify student achievement in relation to the *Senior High ELA Program of Studies*, it is reasonable to expect that the English 30-1 exam focuses explicitly, on measuring students' ability to demonstrate these skills.

The English 30-1 Writing Exam

The writing component of Alberta Education's English 30-1 diploma exam is representative of the current predominant model of large-scale, high-stakes writing assessment. It is administered annually to more than 20,000 students in Alberta. It accounts for twenty-five percent of each student's mark in their English 30-1 course. The completion of this course is a university entrance requirement in Alberta. (For examples of the English 30-1 diploma exam and scoring guides see Appendix 2.)

The writing exam is designed for students to take up to a maximum of three hours to complete. The exam itself is divided into two components. The first component, the

Personal Response to Texts Assignment, is worth forty percent of the exam. The second component of the exam, a *Critical/Analytical Response to Literary Texts Assignment*, is worth sixty percent of the writing exam. The questions for the first component and the second component are always linked thematically. The theme of the exam changes with each administration. Because the exam questions are thematically linked, students are encouraged to read the entire exam over before writing. The *Personal Response to Texts Assignment* is designed to stimulate student thinking for the *Critical/Analytical Response to Literary Texts Assignment*. For this reason students are encouraged to explore the thematic issue in greater depth in the second assignment than they had in the first assignment. Students are told that "time spent in planning may result in better writing" (Alberta Learning, 2005, p. 3).

The exam permits students to respond to the questions from a number of perspectives, either personal, critical, or creative. As well, students are free to express their ideas in any prose form that they deem appropriate to the ideas they wish to express. The exam prompts students both to pay particular attention to the clarity of their communication and to explore their own ideas in their writing.

Personal Response to Texts Assignment

The suggested time for students to complete this assignment is between forty-five and sixty minutes. Before writing, students must read through the print texts and the visual texts provided. These texts are followed by a prompt which places the text into context or which focuses the students' attention to elements of the text that are most relevant to the writing prompt that follows. The prompt in the June 2005 version of the exam reads:

> In the excerpt from *Snow Falling on Cedars*, the writer describes the dual consequences of the islanders' prolonged mutual dependence. In the poem "Casting and Gathering," the speaker observes the tension that exists between opposing perspectives. "Wolf Greeting—Rogue & Pretty Girl—1986," a photograph of two timber wolves, reflects the interplay between dominant and submissive impulses. (Alberta Education, 2005, p. 7)

Following the prompt, student are asked to respond to the following question:

> **What do these texts suggest to you about the ways in which the desire for independence and the need for security shape an individual's identity? Support your idea(s) with reference to one or more of the texts presented and to your previous knowledge and/or experience.** (Alberta Education, 2005, p. 7)

Below the question, is a series of reminders for students:

- select a *prose form* that is appropriate to the ideas you wish to express and that will enable you to effectively communicate to the reader
- discuss ideas and/or impressions that are relevant to this assignment (Alberta Education, 2005, p. 7)

This section of the exam is graded according to two five-point analytic scales. The first scale, *Ideas and Impressions*, is focused on the quality of students' ideas, reflection, and exploration of the topic. It also focuses on how effectively they support these ideas, reflections, and explorations. *Presentation*, the second scale, focuses on:

- the effectiveness of voice and its appropriateness to the context created by the student
- the quality of language and expression
- the appropriateness of development and unifying effect to the prose form.

Markers are prompted to consider the proportion of error to the complexity and length of the response. The scale is somewhat relative, within different contexts certain types of errors will be scored more severely than others.

Critical/Analytical Response to Literary Texts Assignment

The suggested time for this assignment is between one and a half and two hours. Students are provided with a writing prompt:

> Reflect on the ideas and impressions that you discussed in the Personal Response to Texts Assignment concerning the ways in which the desire for independence and the need for security shape an individual's identity. (Alberta Education, 2005, p. 8)

Following the prompt is the question that students are expected to answer:

> **Consider how the desire for independence and the need for security have been reflected and developed in a literary text or texts you have studied. Discuss the idea(s) developed by the text creator(s) about an**

> **individual's attempt to reconcile the desire to act independently with the need for security.** (Alberta Education, 2005, p. 8)

Students are also provided with a series of reminders for planning and writing:

> - You must focus your discussion on a literary text or texts *other than* the texts provided in this examination booklet.
> - When considering the work(s) that you know well, select a literary text or texts meaningful to you and relevant to this assignment. Choose from short stories, novels, plays, screenplays, poetry, films, or other literary texts that you have studied in English Language Arts 30–1.
> - Carefully consider your *controlling idea* or how you will create a strong *unifying effect* in your response.
> - As you develop your ideas, support them with appropriate, relevant, and meaningful examples from literary text(s). (Alberta Education, 2005, p. 8)

Students are provided with ten pages for writing and ten pages for planning.

The assignment is marked using five, five-point analytic scales: a) *Thought and Understanding* is focused on how effectively the students' ideas relate to the assignment and on the quality of the literary interpretations and understandings; b) *Supporting Evidence* is focused on the selection and quality of evidence and on how well the supporting evidence is integrated, synthesized and/or developed to support the student's ideas; c) *Form and Structure* is focused on how well the student's organizational choices result in a coherent, focused, shaped, and concluded discussion and in a unifying effect or a controlling idea that is developed and maintained; d) *Matters of Choice* is focused on how effectively students' create voice through their use of diction, syntax, and other factors; e) *Matters of Correctness* focuses on the student's correct use of sentence construction, usage, grammar, and mechanics. Markers are required to consider the proportion of error to length and complexity when assessing Matters of Correctness.

Summary: The Test and the Construct

To determine what the English 30-1 writing exam values, one must look at the content, the scoring mechanisms and the structure which collectively constitute the exam. An analysis of the content and scoring mechanism reveals the following: The exam values knowledge about language structure—the structure of ideas, of paragraphs, of sentences. The exam also values knowledge about language as a tool through which one communicates ideas. To this end, it values idea formation and support, and it values the creation of appropriate voice. Knowledge about voice is complex requiring knowledge about diction, syntax and punctuation.

An analysis of the exam's structure also reveals the knowledge and skills valued by the exam. Primary among these values is one's ability to generate, organize and effectively present one's ideas within tightly controlled timeframes. As a consequence of this emphasis on time controls, the exam also seems to place a value on one's ability to work effectively under pressure.

It is also important to think of the exam in terms of what it does not value. Given its short timeframes, the exam neither values knowledge about, nor the skill involved in, developing a recursive critical writing process. It is impossible for students to work through an effective recursive writing process while completing two essays in three hours. The exam values a limited form of writing process; in its reminders to students it merely calls for planning, drafting and polishing. The exam ignores substantive revision as an element of writing process. Its scoring criteria, too, do not measure writing process.

The construct being measured by the test is product oriented. While students are told that planning may help improve their writing, they certainly are not assessed on the basis of their planning, nor on any other element of their process. Students are told to focus on materials and ideas that are meaningful to them, however, the narrow focus of the essay questions may limit the student's ability to follow this advice. A key element of the construct is the student's ability to interpret literature and the student's ability to provide meaningful evidence to support the interpretation. The most highly emphasized elements of the construct were the students' ability to develop quality ideas, to support those ideas, and to organize those ideas in such a way that they create a unifying effect. A final element of the construct is the students' ability to polish their writing to enhance its quality.

Given the tight time frame for the examination (the student must write two compositions in three hours), students' ability to formulate ideas quickly, and their ability to organize, shape, and polish their writing under tight deadlines is also a skill that would certainly affect performance and as such must be considered part of the construct being measured.

Comparing Constructs: The Test and the Theory

The construct measured by the writing component of the English 30-1 diploma exam is largely reflective of the construct "good writing" as it appears in the literature and in the *Alberta's Senior High ELA Program of Studies.* The exam focuses on students' ability to develop original ideas, create a strong voice, and adhere to an effective organization. Within this context it also values effective use of conventions.

The constructs differ significantly, however, when it comes to writing process. The diploma exam devalues writing process while both professional writers and acclaimed educators place a premium on it. The single-sitting, time constrained, impromptu structure of the exam devalues process in a number of ways: a) it does not provide adequate time for students to engage in a meaningful two stage—creative, and critical—process; b) the combination of time constraints and a high-risk context encourages "safe writing" rather than risk-taking and exploratory thinking. This combination of constraints also encourages single drafts; c) The impromptu nature of the exam virtually eliminates any possibility for students to create a text that is personally meaningful to them; rather, it requires students to quickly formulate ideas that correspond to the question being asked; d) The rigid time-frame does not allow for students to adequately revise or reconceptualize their writing.

The exam construct also differs significantly from the construct contained in the curriculum on the basis of its simplistic treatment of the issue of metacognition. The Grade 12 program of studies places a strong and consistent emphasis on metacognitive skill development in student writers. It requires students to reflect on the choices they make as writers, to critically examine those choices and their impact on the text, and to develop strategies with which to improve their choices and their use of strategies. The diploma exam asks students to reflect on their choice of text when responding to the major writing question. However, their metacognitive reflection is not included in the portion of the student response that is assessed. While reflecting on choice of text might be helpful, the program of studies calls for a far more significant and extended use of metacognition in the context of student writing; it focuses on the use of metacognitive skills to enhance the processes through which students write.

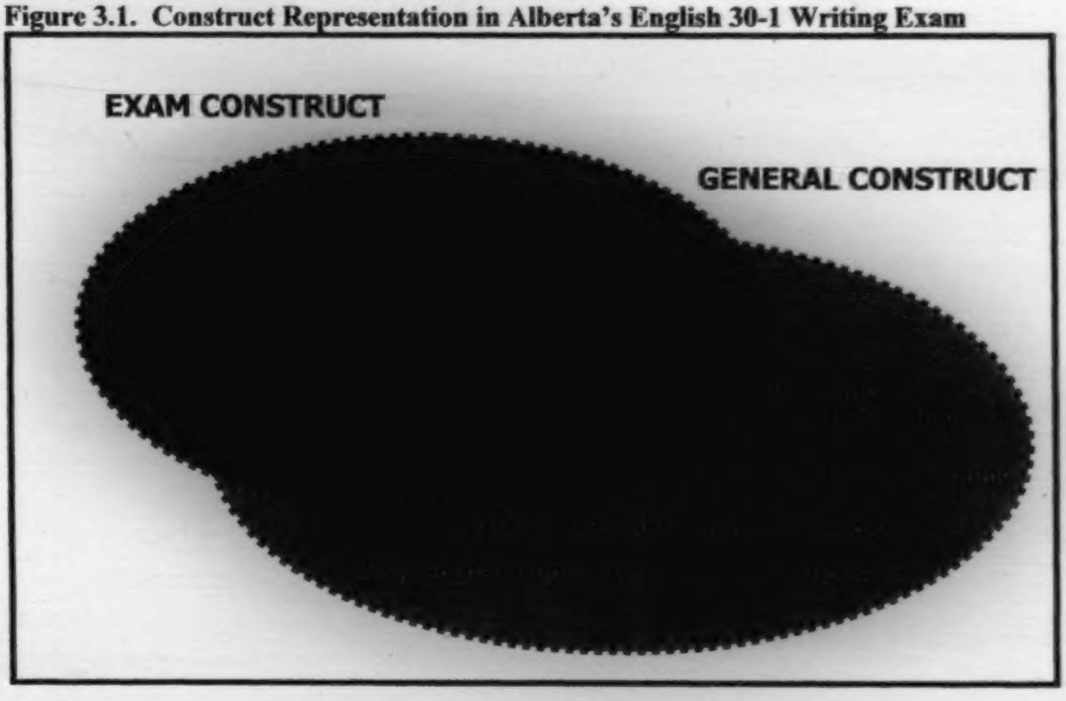
Construct under-representation and irrelevant-variance. The concepts of construct under-representation and construct irrelevant-variance were first introduced by Cook and Campbell (1979). In describing their list of threats to construct validity they write:

> They all have to do either with the operations failing to incorporate all the dimensions of the construct, which we might call "construct underrepresentation," or with the operations containing dimensions that are irrelevant to the target constructs, which we might call "surplus construct irrelevancies." The list concentrates mostly on the fit between constructs and the way that the research problem is conceptualized, and devotes less attention to generalizing across constructs. Getting the initial question "right" is not as important a construct validity issue as getting one's operations to reflect one's research constructs. (p. 64)

Essentially, they argue that the main concern regarding construct validity is that the instrument measures what it claims to be measuring, rather than that the construct being measured in fact reflects general understandings. Historically, this perspective on construct underrepresentation and construct irrelevant variance has received the focus in construct validity research. However, the first item on Cook and Campbell's list of potential sources of construct underrepresentation and construct irrelevant variance focuses on "inadequate preoperational explication of constructs" (p. 64). This issue revolves around the failure to clearly define or delineate one construct from another. As a consequence of this failure, results derived from a measurement tool may be based upon incomplete understandings of the construct which in turn raises questions about inferences and scores derived from the instrument. The *Standards* (AERA, APA & NCME, 1999) point out that "nearly all tests leave out elements that some potential users believe should be measured and include some elements that some potential users consider inappropriate" (p. 10). The burden, then, for those critical of an assessment's construct is to demonstrate why one set of elements should be included in the test's construct. The critique expressed in this study, focuses on two issues:

First, Alberta's English 30-1 diploma exam is designed to certify student achievement in relation to the Alberta English 30-1 program of studies. Second, the diploma exam is designed to certify student writing ability. However, Alberta's English 30-1 diploma exam fails to adequately capture the construct "good writing" as described in both Alberta's *Senior High ELA Program of Studies* and within generally held understandings of this construct (as described in Chapter 2 of this study). Given the exam's design and grading system, it can be argued that important general construct elements are neglected by the exam's construct and that important extraneous construct elements are introduced into the exam's construct. These areas of neglected and extraneous construct elements are depicted in Figure 3.1.

Figure 3.1. Construct Representation in Alberta's English 30-1 Writing Exam



Skills or attributes contained in the oval titled "General Construct" represent generally agreed upon elements of the construct as described within the literature in Chapter 2 of this study. Skills or attributes contained in the oval titled "Exam Construct" represent elements of the construct the exam's scoring guide, structure, and writing context suggest the exam is measuring. Items in the "General Construct" oval which are not contained within the overlapping area between the two ovals are construct elements I would suggest are being ignored by the diploma exam, while items in the "Exam Construct" oval that are not in this overlapping area are construct elements measured by the exam which are extraneous to the generally held understanding of the construct.

The purpose of this study is not to examine whether or not these problems with construct representation are adversely affecting students' scores on this exam, or whether these flaws are negatively affecting some students in unintended ways (though a study of this nature would be valuable). Rather, the focus of this study is on whether or not these perceived flaws in representation are impacting the pedagogical choices teachers are making when preparing students for this exam. This study also focuses on what messages students take away from their experiences of learning to write within the context of this diploma exam.

I argue elsewhere (Slomp, 2005) that government-mandated, high-stakes assessments are the clearest and most public statement regarding what skills and knowledge are valued within a given system of education. My focus in this study, therefore is less on inferences drawn from the exam's scores, and more on individuals' transactions with the exam; it focuses on what students and teachers learn from this exam regarding which skill-sets and knowledge-domains are being emphasized by Alberta's English 30-1 assessment program. It is the consequences of these transactions on pedagogical choices and student learning that are the primary concern of this study. As such, while this study draws on important concepts from validity theory it should not be considered a validity study in the traditional sense; it is a study about composition pedagogy within the context of a high-stakes, standardized writing assessment.

CHAPTER 4: MIXED METHODS RESEARCH

We work in our own darkness a great deal
with little real knowledge of what we are doing.
–John Steinbeck

Technique alone is never enough. You have to have passion.
Technique alone is just an embroidered pot holder.
- Raymond Chandler

I don't know about method.
The what is so much more important than the how.
-Ezra Pound

In Chapter Two, I discuss in some detail the issue of consequential validity. Questions regarding the impact of assessment on teaching and learning focus on important aspects of consequential validity. So, while this study focuses on issues of pedagogy and learning, it also addresses important questions regarding the consequences of Alberta Education's English 30-1 diploma exam on composition pedagogy in Alberta's English 30-1 classes.

In the late 1990's a series of articles (Linn, 1998; Lane, Park & Stone, 1998; Green, 1998; Shepard, 1997; Popham, 1997, 1999; Yen, 1998; and Moss, 1998) which focused on consequential validity issues provided a series of recommendations regarding the design of studies which focus on the consequences of assessment. According to these scholars, studies of test use consequences should:

(a) investigate the "actual discourse and actions that occur around products and practices of testing;" (Moss 1998, p 7)

(b) corroborate data collected from multiple sources (i.e., teachers, students, and administrators);

(c) develop both comprehensive sources of direct evidence collected in classrooms as well as more global sources of evidence such as surveys;

(d) be highly contextualized, intensive, and sustained.

Collectively these authors argue that, given the types of questions asked and the information collected, research into the consequences of assessment would best be conducted within a mixed method design. Greene (2005) supports these recommendations and suggests that in the context of program assessment mixed methods approaches to research are essential. She observes:

> Again, evaluation approaches and methods differ by their attention to one particular context, or case, or their emphasis on cross-context regularities. Today's political pressures are barely cognizant that educational settings are 'complex and changing networks of human interaction' that fundamentally cannot be controlled (Berliner, 2002 p. 19), nor that one educational context can be vastly different from another in ways that significantly interact with the quality and effectiveness of teaching and learning. With a mixed methods approach, cross-context patterns of regularity *and* within-site contextual complexity are *both* respected and engaged. The strategy is one of dialogue and conversation between these two perspectives. (Green, 2005, p. 210)

Based on this argument, and following the above recommendations, this study has been designed as a mixed methods study, one which incorporates both case study and survey methods.

Mixed Methods Research

Tashakkori & Teddlie (2003) claim that a multiple-case-study-plus-survey model is a classic form of mixed methods research design. Commenting on its strengths, they suggest that, "[o]ne method gives greater depth, while the other gives greater breadth; hopefully, together they may give results from which one can make better . . . inferences' (p. 16).

Increasingly, researchers are advocating for multi-method approaches to social sciences research (Creswell 2003; Johnson & Christensen 2004; Palys 1997; Tashakkori & Teddlie, 1998, 2003). Tashakkori & Teddlie (2003) argue that the multi-method approach has been used since the early 1900s. During the 1980s and 90s, however, "paradigm purists" argued that qualitative and quantitative methodologies were incompatible with one another. They further argued that because of these differences in underlying philosophies both methods could not be successfully combined. Advocates of mixed methods, however, point at the one hundred year history of mixed methods research to refute that claim (Tashakkori & Teddlie, 2003). Their position is further bolstered by Denzin and Lincoln (1998) who claim that neither quantitative nor qualitative methods are exclusive to any one paradigm, rather they suggest that a range of methods can be used appropriately within a variety of paradigms.

More recently, however, Denzin and Lincoln (2005) express some reservations about mixed methods designs. They suggest that historically, mixed methods researchers have taken “qualitative methods out of their natural home, which is within the critical interpretive framework” (p 9). They also observe that most mixed methods studies place qualitative methods in a subordinate position to quantitative methods; within these designs qualitative methods are seen to be exploratory while quantitative methods are seen to be confirmatory. They further critique the quantitative elements of the mixed methods research for their tendency to strip away context.

While Denzin and Lincoln’s critique is important, it might be a touch premature; qualitative perspectives on mixed methods research are now emerging. Green (2005), for example, claims that mixed methods research stem from a qualitative mentality, one that values diversity of method and which recognizes the importance of contextual diversity. She states, “In short, a mixed method way of thinking is itself anchored in values of toleration, acceptance and respect—of multiplism and difference” (p. 208). The goal of mixed methods research she argues is to develop complex understandings through the application of multiple lenses, perspectives, and stances while also challenging simplistic answers to complex questions. Mason (2006) expands upon this argument suggesting that mixed methods research can enable researchers to see outside the box. She observes that conventional wisdom suggests that our theoretical orientations shape our methodological practice, which in turn shapes the kinds of questions we can ask and answer. She then argues that the opposite is also true: “Our ways of seeing, of framing questions, are strongly influenced by the methods we have at our disposal, because the way we see shapes what we can see, and what we think we can ask” (Mason, 2006, p. 13). She argues then, that mixed methods research, stemming from a qualitative way of thinking, enables researchers to explore through more diverse means, the complexities of social phenomena.

This study approaches mixed methods research from a qualitative perspective. The case study component of this work is understood to be fundamental to the study as a whole. Its focus on exploring the range of contextual issues that teachers are facing when preparing students for the English 30-1 exam is an important concern for this study. Rather than taking qualitative methods out of their home, this study applies quantitative methods within a critical interpretive framework. Survey questions focus on developing a picture of student context and on exploring students' critical perspectives on their English 30-1 experiences while at same time looking at trends in relation to understandings and context across a large number of students. Table 4.2 (at end of this chapter) clearly demonstrates the importance that this study places on context: While this study is interested in the English 30-1 diploma exam's influence on teaching and learning, it focuses significant attention on students' and teachers' contexts, and the impact that their contexts exert on either enhancing or mitigating the exam's influence.

Advantages of Mixed Methods Research

In spite of Denzin and Lincoln's (2005) concerns, many qualitative researchers are adopting mixed methods designs. Denzin and Lincoln (2005) offer two explanations for this development. On the one hand, they suggest, this movement may simply be a response to the political realities of the time for social sciences researchers; or, it might suggest a "backlash against the perceived excesses of postmodernism" (p 10). While these reasons may certainly have some merit, one might also suggest that a movement towards mixed methods designs merely recognize the strengths of such designs. Johnson & Christensen (2004) suggest a number of advantages for conducting mixed research. Mixed methods research:

- enables the use of words, pictures, and narrative to add meaning to and to provide context for quantitative data;
- enables the use of quantitative data to add precision to words, pictures, and narrative;
- draws upon the strengths of both quantitative and qualitative research methods;
- enables researchers to generate and explore a grounded theory;
- enables the researcher to answer a broader and more complete range of research questions because the researcher is not confined to any single method or approach;
- provides stronger evidence for a conclusion through convergence and corroboration of findings;
- enhances insights and understandings that might be missed when only a single method is used;
- produces more complete knowledge which can be used to inform both theory and practice. (p. 414)

In their historical overview of mixed method research, Tashakkori & Teddlie (2003)—reflecting a positivist perspective—suggest that the three most important qualities of these research designs are their ability to both generate and test theory, their ability to utilize complementary strengths while ensuring against nonoverlapping weaknesses, and their ability to use triangulation or corroboration of data to enhance the researcher's confidence in the study. Mason (2006), arguing from a qualitative perspective, suggests three additional strengths of mixed-methods designs: a) they enable researchers to explore social phenomena through a range of methods and lenses, thereby developing a richer sense of the complexities involved; b) they enable researchers to explore "how social experience and 'real lives' are simultaneously or connectedly 'big and little', global and local, public and private" (p. 15); and c) they enable researchers to "enhance and extend the logic of qualitative explanation" (p. 16).

Drawing on the strengths of quantitative and qualitative designs. Essentially, Mason's argument is constructed upon the idea that if used properly, mixed methods research takes advantage of the strengths of both quantitative and qualitative research methods. Johnson & Christensen (2004) list a number of strengths for each design framework (see figure 4.1).

My use of quantitative data enables me to explore the experiences of a large number of people; it enables me to assess whether or not themes emerging from case study explorations resonate more broadly with the experiences of others. My use of qualitative data enables me to explore contextual details in far greater depth than my survey will allow, it enables me to understand the phenomena from the point of view of my participants, and it enables me to form contextually rich understandings which I will be able to explore in future work. Combining both case-study and survey methods will enable me to explore both the micro (rich individual contexts and experiences) along side the macro (patterns of student experiences and critical understandings of approaches to writing).

Table 4.1. Strengths of Quantitative and Qualitative Research Methods

Quantitative methods can:	Qualitative methods can:
• be used to test and validate already existing theories	• be used to develop tentative explanatory theories about a phenomenon
• be used to test preconceived hypotheses	• be responsive to local situations, conditions, and stakeholder needs
• generate findings that are generalizable	• enable participants to explore personal experiences of phenomena
• allow for quantitative predictions	• lend themselves to exploring how and why phenomena occur
• enable researchers to study variables in isolation from one another	• be useful for describing complex phenomena
• be used to collect data relatively quickly	• generate data based on participants' own categories of meaning
• be used to analyse data relatively quickly	• be used to study dynamic processes
• be used to generate results that are relatively independent from the researcher	• describe in rich detail phenomena as they are situated and embedded in local contexts
• generate results that often possess higher degrees of credibility with people in positions of power	• determine how participants interpret constructs
• be useful in studying large numbers of people	• be useful for studying a limited number of cases in depth

Limitations of Mixed Methods Research

While emphasizing the strengths of mixed method's research Johnson & Christensen (2004) also suggest several weaknesses to this method. These include the difficulty involved in carrying out the research in terms of work involved, time required, expenses incurred, and knowledge required to collect two types of data. They also point out that methodologists in the field of mixed research designs have yet to resolve outstanding issues regarding paradigm mixing, and coherently interpreting multiple sets of data. While these issues certainly are worth considering, I think the strengths involved in using this approach outweigh these concerns.

Mixed Data Collection Methods

Based on these considerations of strengths and limitations, this study was designed according to a mixed methods design. The case-study component of this study was conducted with three English 30-1 teachers in a rural, a rural/urban, and an urban school. Each case study utilized three methods of data collection: Several weeks of direct observation of classroom activities; interviews with each teacher, interviews with ten students; and the collection and analysis of all writing assignments (teacher handouts, not student work) and pertinent classroom documents. A more detailed description of the data collection and analysis methods for this component of the study will be discussed in Chapter 5. The survey component of this research focuses on developing a broader picture of students' experiences with learning to write in English 30-1. One hundred and thirty-two students from seven schools participated in the survey component of this research. A more complete description of the methods involved in the survey component of this study will be discussed in Chapter 6.

While the methods used in this study are diverse, the broad spectrum data collection they enable provide for a significantly rich and multi-faceted picture of the diploma exam's influence on teaching and learning, and on the contextual variables that either enhance or reduce the exam's influence. The material collected through each method of data collection is not treated in this study as an individual data set; rather, this data is understood to be intimately linked to, and informed by, data collected through other means. The inter-related nature of data collection within this study is represented in Table 4.2.

Table 4.2. Overview of Data Collection Methods

	Questions	Primary Method	Data Collection	Discussed in chapter
What contextual factors might enhance or mitigate this exam's influence on the teaching and learning of composition?	How does the writing component of the English 30-1 diploma exam influence students' beliefs about, understandings of, and approaches to creating effective writing?	Student interviews	Focus on student writing process, their attitudes toward the diploma exam, their interpretations of their teachers' writing assignments, and their experiences with learning to write.	5 & 6
		Surveys	Focus on critical perspectives on diploma exam and classroom assessment constructs, attitudes toward diploma exam, writing practices, classroom writing experiences, and beliefs about writing.	6
		Classroom observation	Focus on writing habits, and classroom writing experiences.	5 & 6
	In what ways does the writing component of the English 30-1 diploma exam influence teachers' pedagogical and classroom assessment choices in their English 30-1 classes?	Teacher interviews	Focus on educational background and experience, beliefs about teaching (in general and relation to writing), personal writing experiences/perspectives, attitudes of school leadership and community, and perspectives on the diploma exam.	5
		Classroom Observation	Focus on pedagogical choices, organizational structures, student – teacher relationships, and use of class time for writing assignments.	5 & 6
		Document Analysis	Focus on pedagogical choices; classroom assessment emphasis and constructs; and type, range, and frequency of writing assignments.	5

Table 4.2 points to the foci of the following two chapters. They focus first on teachers' and then on students' transactions with the diploma exam. Embedded within rich contextual data, these chapters explore the impact of these transactions on teaching and learning.

CHAPTER 5: THREE TEACHERS

Writers, if they are worthy of that jealous designation,
do not write for other writers. They write to give reality to experience.
- Archibald MacLeish

All history becomes subjective; in other words
there is properly no history, only biography.
- Ralph Waldo Emerson

Koretz and Hamilton (2006) argue that if one wants to understand what traits or construct elements a test is primarily assessing one should explore teachers' analysis of the test. After all, they observe, it is part of the teachers' job to prepare students for external assessments. Given this responsibility, teachers have a natural incentive to deconstruct external assessments to determine what they are primarily measuring.

Having made this claim, Koretz and Hamilton (2006) ask, what research methods might be suitable for collecting this type of information. The case study approach chosen for this study is one effect method through which this information can be collected.

This chapter begins with a discussion of the case study methods used in this study and then turns its attention to describing and analyzing three teachers' transactions with Alberta's English 30-1 diploma exam.

Case Study Methods

The case study element of this research is based upon a multiple case study design. The study is instrumental in focus. Instrumentally focused, multiple case studies are interested in specific issues or phenomenon rather than on the intrinsic interest in the specific cases themselves. Based on this set of interests, Stake (2006) recommends that multiple case study research participants be selected on the basis of two criteria: Diversity and opportunity to learn, with the latter criteria—the potential richness of information and ease of access they offer—being the primary consideration. While the case is considered to be a secondary concern for instrumental case study research, it remains the primary focus of the study during the data collection phase and during the initial data analysis phase. During these phases of the research, the research focus rests primarily on the desire to first collect and then develop a rich description of each case. This rich description is necessitated by a qualitative sensibility, one which recognizes the socially, culturally, situationally, and contextually mediated nature of our perceptions of reality. Rich descriptions from multiple viewpoints can help us to understand the phenomena under investigation more completely. Stake (2006) concludes:

> Thus the methods for case study work actually used are to learn enough about the case to encapsulate complex meanings into a finite report but to describe the case in sufficient descriptive narrative so that readers can experience these happenings vicariously and draw their own conclusions. (p. 240)

Further in this chapter, three cases are presented each with its own detailed narrative. I have tried to keep these narratives as tight as possible; their length reflects an appreciation of Stake's advice while simultaneously it invites the reader to engage in his or her own transaction with the text.

Stake (2006) warns that only after the rich descriptive work has been completed can the researcher then focus on conducting an analysis across cases. The final section of this chapter contains an analysis across cases which explores common experiences, contexts and issues.

Selecting Participants

My search for participants began by consulting leaders in English Education in Alberta and in the area in which I lived. I discussed my research with the provincial executive of the Alberta Teachers' Association's English Language Arts Council, with district superintendents, with local principals, and with graduate students at the University of Alberta, requesting recommendations from them regarding potential teacher participants. Six potential participants were contacted. Three were selected.

The participants selected represent diversity in a number of contextual factors:

- their number of years teaching,
- their number of years teaching English language arts,
- their years of involvement with the diploma exam marking team,
- their level of involvement with the diploma exam marking team (marker, table leader, or standard setter),
- their gender,
- their school context,
    - urban or rural,
    - large or small,
    - semestered courses or whole year courses,

- public school, Catholic school, or independent school

While this predetermined diversity does not guarantee that the participants involved are representative of a larger sample of teachers, Stake argues that representativeness is not a quality that is often strived for in case study research. The diversity in terms of participants' backgrounds and contexts does, however, allow for discussion related to how contexts might have an impact on approaches to pedagogy.

While each participant profile will be developed in depth later in this chapter, a few lines of introduction seem appropriate at this point: Anne teaches in a rural junior/senior high school. She is in the eighth year of her teaching career. Brian teaches in a urban/rural K-12 school. He is in the eight year of his teaching career. Heather is the English language arts department head in a large high school which is located in an urban setting. She is in the fifteenth year of her teaching career. Pseudonyms are used throughout this study when referring to all participants, both teachers and students.

Anne, Brian and Heather each provide significant opportunity to learn: They are experienced teachers of writing who are enthusiastic about participating in the research; they were willing to allow extensive observations in their classrooms; they readily provided me with their teaching documents; and they willingly participated in three lengthy interviews.

Data Collection

Data for this element of the research were collected through interviews with each teacher participant; through the collection of writing assignments used by each teacher participant in his or her English 30-1 course; through classroom observations of each participant's English 30-1 class; and through interviews with three or four students from each participant's class.

Teacher Interviews

Each teacher participated in three interviews, two individual and one group interview. The first interview occurred at the beginning of the research project, during the month of February 2005. This interview was semi-structured and lasted between forty-five minutes and one and a half hours. This interview focused on perceptions of themselves as teachers of writing, a description of their approach to teaching writing, experiences which have shaped their approach to teaching writing, their school context and its impact on their approach to teaching, their perceptions regarding the diploma exam, and the exam's impact on their approach to teaching writing. The purpose of this interview was to begin to collect information regarding individual backgrounds, contexts, and perspectives.

The second interview occurred toward the end of the school year, in late May or early June 2005. It occurred after classroom observations and the interviews with students had been completed. This interview was less structured, focusing instead on questions that presented themselves during the preliminary analysis of the first round of interview data, during student interviews, or during classroom observations. A significant number of questions were designed to help the researcher confirm or challenge his perceptions of the teacher, his/her practice, and his/her beliefs about teaching, assessment, or writing.

The third was a group interview which occurred early in the fall during the following school year (October, 2005). This interview was free ranging, mostly directed by the teacher participants who collectively explored their perspectives on a range of issues important to them. They discussed societal issues and their impact on teaching writing, issues related to student writing, the differences between their own writing processes and the processes they ask of their students, the value of writing across the curriculum, and their perspective on the diploma exam. The conversation saw frequent moments of agreement and extension but also moments where the three teachers disagreed with one another. While this interview touched on ideas discussed in the previous two interviews, the discussions between teachers help to enrich and expand on the ideas expressed in the previous interviews.

Interview data analysis. Each interview was recorded and fully transcribed. Transcriptions were submitted to the teachers for review and comment. Each teacher approved the transcripts of his or her interview. Total length of transcripts for all three sets of interviews came to 88 pages of single spaced text or 51, 650 words. Each interview was analyzed first with a focus on general themes: Background/context, pedagogical issues, assessment issues. Some aspects of the teachers' commentary were reflected more than one general theme; for example comments which discussed the relationship between both assessment and pedagogy were included into both the pedagogical and the assessment themes. A second finer layer of analysis, which looked through each general theme focusing on developing sub-themes then took place. Within the category "Background/Context" seven themes sub-themes were identified:

- Motivation
- School context
- Personal writing (style and process)
- Previous educational experiences
- Attitude toward standardized assessment
- Professional development experiences
- Perspectives on writing pedagogy

Within the category "Pedagogical Issues" six sub-themes were identified:

- Writing process, perspectives and practice
- Personal writing process and its impact on pedagogy
- Pedagogical tools/focus
- Planning for instruction
- Perspectives on students
- Issues of importance

Within the category "Assessment Issues" there sub-themes were identified:

- Perspectives on the English 30-1 diploma exam
- Perspectives on the English 30-1 diploma exam's construct
- Perspectives on the English 30-1 diploma exam and its influence on pedagogy

This thematic scheme provided a structure within which to further refine the analysis of the interview data. Within these categories, teachers' comments were grouped around similar sub-themes. For example, Brian commented across the three interviews on the issue of student use of time for completing writing assignments. These comments were first categorized under Pedagogical issues, from there under issues of importance, and then they were further grouped within that category around the issue of time. Following an approach used by Rex and Nelson (2004), I then created a pastiche which brought together the teachers' comments around each of these specific categorizations. The resulting pastiche regarding Brian's comments about time was finalized as follows:

> *Partly I think [students follow a limited writing process] because they tend to wait to the last minute so they don't allow themselves time to rework things, to look at it and ask does my organization make sense? Or, am I putting too much emphasis on this or not enough emphasis on that? I don't think that is always the case, there are some kids that work very hard at trying to change it, trying to make it better, but I do think that is one of the reasons. That is why I tend to give shorter deadlines now. . . . I have learned, you can give students three weeks to do an essay and chances are most of them are punching it out the last night. So I think shorter deadlines work. . . . I think long due dates are good for some kids because it gives them—I know for myself when I was in university you'd have an assignment that was due a month down the road and it's not that you right away start writing but you were thinking about it—but I find that for high school kids they don't do that and they tend to leave things. So I try to force the process by just giving short due dates.*

This pastiche is a compilation of three separate comments on the issue of time in relation to student writing. Collectively this pastiche captures more completely Brian's perspective on student use of time than the individual comments would have. Rex and Nelson (2004) contend that no representation of an individual is ever complete and that a pastiche of this nature also does not capture perspectives or individuals completely but that such a pastiche can provide a means to more eloquently and purposefully represent an individual's perspective while maintaining that individual's voice. Out of respect for the individuals participating in this research, each pastiche is punctuated to reflect the voice of the speaker as accurately as possible, therefore the punctuation reflects conventions of spoken English rather than written English. Also, each pastiche has been edited with a focus on reflecting the intent of the speaker rather than capturing his or her comments verbatim: pauses, stutters, repetitions, and other features of spoken language have been edited out of these pastiches.

Discussion of teacher identity, practice, and perspective in this research is built around these pastiches.

Student Interviews

In addition to interviewing the three teacher participants, I also interviewed a total of ten of their students: three each of Brian and Anne's students and four of Heather's students. During the second teacher interview, each teacher was asked to identify three students who represented a range of writing abilities—one weak writer, one average writer, and one strong writer—each of whom might be willing to participate in an interview. The students who were identified agreed to be interviewed. Given the requirements of the research agreement with Anne's school board (who did not want research taking place during the month in which exams were taking place), her students were interviewed prior to writing the diploma exam. At the time they were interviewed, however, they were quite familiar with the writing context for the diploma exam because they had written a number of practice exams during the school year. Brian and Heather's students were interviewed on the day after they had written the diploma exam. Each interview took place in a private setting within the school (the school counselor's office or an empty classroom) and took between half an hour and an hour to complete.

The student interviews were semi-structured. They focused on the students' perceptions of themselves as writers, their perceptions of their English 30-1 teacher as a teacher of writing, their writing processes, and their experience writing the exam (or exam preparation writing).

Interview data analysis. Each student interview was fully transcribed prior to analysis. Transcripts for all student interviews totaled 56 pages. Interview data was used both for the case study and for the survey elements of this study. The case study element focused mostly on teachers and teaching, and the survey element of the study focused mostly on the student and the learning of writing. (I will report on the analysis of survey related materials in Chapter Six.) Case study analysis of student interview data focused on the student's perceptions of the teacher, the teacher's teaching style, and the teacher's focus when marking writing assignments. Each student's perspective on the teacher and his/her marking was included in the case study profile of each teacher.

Classroom Observations

The second method of data collection involved classroom observations. The duration of observations was different for each teacher and depended on the unit he or she had chosen for me to observe and whether or not the school ran on a semestered or full-year system. Observations took place during May and early June, 2005. Anne's English 30-1 course was semestered—I observed her class every morning for two and a half weeks. Brian and Heather's classes were full year. I observed Brian's class for three and a half weeks and Heather's for four weeks. I also came to Heather's three exam preparation workshops, offered early in the morning before school began.

Classroom observations focused on teaching style, pedagogical stance, student and teacher's comments regarding the diploma exam, and classroom environment. Field notes capturing these observations were taken.

Analysis of observation data. The purpose of collecting the observation data was to enable the researcher to better understand the teacher, the teaching environment, and the teaching practice. The interview data provided the researcher with the teachers' construction of him/her self as a teacher; the classroom observations enabled the researcher to also construct a representation of the teacher based on direct observations. These constructions were compared to one another to identify coherence and dissonance. Aspects of apparent dissonance were then discussed in follow up interviews. For example, Anne sharply criticized the diploma exam constructs during the first interview, but I noticed during observations that she was quite focused on the exam in her teaching. The contradiction implicit in these differences puzzled me, so during the second interview I asked her about the apparent contradiction. She responded as follows:

> *It is a complete paradox; You can't do that right. You can't say the exam is not a fair assessment but I am going to use it anyway. But that is what*

> *we do. . . . Because it is really hard to reconcile those two things. You say well, the expectations of the administration, and of the parents, and of the students themselves, is that you prepare me for the exam, ok fine, I can do that but personally I don't feel that this exam is a fair assessment. But that is the expectation, so then I have to balance these two, and wrestle with these two in the classroom, and say okay, I am going to let the exam go a little bit and we are going to do something wild and creative and have a little fun with this piece of literature rather than focus exclusively on the exam. It is really hard to reconcile those things together. . . . Yeah, valid or not, the darn thing still exists, and students are going to have to write it. I mean the only way you could eliminate the tension is to eliminate the exam, and that is not going to happen, that is not going to happen.*

This process enabled the researcher to both develop a more complete understanding of the complex matrix of variables that influence teaching and learning within a high-stakes testing environment and to complete case study profiles which more completely reflect the teachers' views and practices.

The analysis conducted with the interview data provided the foundation for the analysis of the observation data. Observations were linked to themes developed in the interview analysis and were then used to support or enhance case study profiles.

Document Analysis

Teacher participants were asked to provide a copy of each writing assignment or teaching document related to writing that they used with their English 30-1 class. Heather provided 125 handouts which included assignments, writing support materials, marking guides and course or unit outlines. Anne provided 44 complete lesson plans which included assignments, support material, marking guides, unit outlines. She also provided two, two-inch binders which contained her diploma exam preparation materials. Brian provided 26 writing assignments, which included midterm exams, marking guides, end-of-unit assignments, and major projects. While these materials do not represent the complete number of assignments and handouts given by each teacher in English 30-1 class, they do provide significant detail regarding the range and the focus of each teacher's approach to teaching writing.

Analysis of Documents. Each set of teacher documents was analyzed in terms of the constructs the assignments were attempting to measure. Specifically, analysis focused on whether or not the assignments were focused on the construct being measured by the diploma exam. Important elements of the exam construct that were focused on included:

- Was the assignment designed as a response to literature or not?
    - If so, was the assignment designed as either a personal or critical response to literature?
- Did the assignment specify a genre in which students must respond? If so, what genre?
- What time lines were included in the assignment?

- What marking criteria were being used to judge student performance on the assignment?

Handouts were also analyzed in relation to each teacher's comments regarding his or her teaching practice. For example, Brian commented often in the interviews on the issue of time. He suggested that because students did not use their time effectively, waiting often until the last day to work on assignments, he began shortening due dates on his writing assignments. When analyzing Brian's writing assignments, I focused on comments regarding the use of time to see what other strategies he might have used to encourage students to use their time more effectively.

Information generated through document analysis was incorporated into case study profiles to help provide a more complete picture of each teacher's teaching practice.

Case Study Profiles

In his chapter on case study research in Denzin and Lincoln's (2005) *Handbook of Qualitative Research*, Stake (2005) argues that triangulation is an essential element in collective case study research. In case study research, triangulation is commonly achieved through the use of multiple procedures for collecting and analyzing data, through redundancy in data collection, and through use of "multiple perceptions to clarify meaning" (p. 452). This study adheres to the principles of triangulation by collecting data through multiple methods and by comparing and contrasting multiple perspectives, both within and across cases. Each case study profile draws on this range of data in an attempt to present as complete a picture, incorporating the range of perspectives (researcher, teacher, student) of each teacher and his/her relationship to the diploma exam as possible.

Once completed, each profile was submitted to the teacher participant for comment and feedback. The profiles which appear below have been read by the teacher participants and have been approved by each of them. Each teacher commented that he or she felt that his or her profile reflected accurately his or her own perception of him/herself and his or her teaching practice.

The profiles have been presented in alphabetical order according to pseudonym, though profiles can be read in any order the reader chooses.

Anne's Profile

Anne has been teaching English language arts for eight years. She has completed a BA in English and French, and a BEd in English. Prior to teaching she worked as a journalist.

Anne teaches English 30-1 in an overcrowded classroom. Her desk is at the front of the room. In the back of the room hangs a large television set, and in the back corner stands a bookshelf containing a range of genres, from classic novels, to graphic novels, to popular fiction and magazines. Students are given free access to the bookshelf. Anne's room has desks for 34 students. These desks are packed together in tight rows facing the classroom's center aisle, on each side of the aisle the rows are three desks deep. Twenty-eight students are enrolled in English 30-1. They have little room to move around and Anne has difficulty getting to the students who are on the perimeter of the room. Yet, given the atmosphere of respect that pervades the room, Anne does not have problems with classroom management. Student representing assignments (visual representations of learning/thinking) are posted on the walls around the room.

Anne teaches in a rural school in a small town approximately 40 minutes from a major city. Most of the school's students are members of farming families. Anne's school serves approximately 355 students from grades 7 to 12. The school's website boasts a strong academic program, citing student performance at or above the provincial achievement levels (average scores on achievement test and diploma exams) to support this claim. The school offers students many opportunities to engage in career-related programs of studies: advanced levels in areas of communication, technology, design studies, construction technology, fabrication studies, mechanics, energy, mines, food studies, fashion studies, and cosmetology.

Anne believes that teaching in a farming community—especially in an era where the family farm is slowly fading into memory—shapes the educational focus of students and parents. As a teacher she feels a certain pressure and urgency from the community to ensure that her students are prepared for a new future. She states:

> *[My school is in] a small town, a small...farming community. These kids know that if they don't get out of this school to go to post secondary of some kind, any kind, they'll be working at the IGA for the rest of their lives or chucking bales on the farm. And they just can't fathom that. That's just not an option for them. So there's so much [riding] on [their school and exam performance]. And also for the parents—why would*

> *anybody choose farming in this day and age? When you die, you bought the farm, how bad is that?—"This is your opportunity kid, seize it now or you're stuck with me for the rest of your life." So I think there is a lot of parental pressure there.*

Anne and her students, in conversations with me, focus on the pressures of schooling and on the expectations to perform. Interviews with several of Anne's students support her understanding of the pressure in their lives; they talk about the importance of doing well so that they can open opportunities for the future.

Other factors seem to intensify the pressure surrounding the English 30-1 course in Anne's school. English 30-1 is a semestered course in Anne's school, running during the second semester. Because summer school courses are only offered in a city half an hour away, students who perform poorly in English 30-1 during Anne's class are more limited in their ability to retake the course than a student studying in a major city.

In addition to the pressures placed on students, Anne notes that significant pressure from parents and school administration members is placed on teachers. This pressure seems to stem in part from the culture of her school community (as evidenced by the school website's emphasis on achievement test results) and from an experience during Anne's first year of teaching English 30-1. That year her students scored on average 20 percent lower on their English 30-1 diploma exam than they did on the classroom awarded mark in English 30-1. Parents were very unhappy with this result. At the time, her principal publicly supported her. However, the new principal who came to the school the following fall made it clear to her that in the future 20 percent differences between school and exam scores would not be acceptable. Since then, Anne has been acutely aware of the pressure to ensure that her school and exam scores are satisfactorily aligned with one another. She reports that unofficially her school administrators expect to see no more than a five percent difference between school scores and exam scores.

In spite of these pressures, Anne enjoys teaching: she enjoys the relationships she is able to develop and maintain with her students, especially her senior high students; she enjoys challenging her students to broaden their thinking and their experiential base; and she enjoys capturing her student's interests. Anne reflects on her pedagogical stance:

> *I teach junior high and senior high at this point. So in a high school setting I try to operate on the principle of mutual respect,. . . . I am fairly laid back in the classroom. I like to be informal with them. I don't make anybody raise their hands, I enjoy a lively discussion. . . . The kids yell at me because I use big words but I think it is important for them to expand their vocabulary and realize that yes people actually do use these words. I really enjoy the part of teaching that helps me to stay young and involved with them, it is a mental thing for me. I enjoy the mental interactions, the discussions, those kinds of things. I enjoy inspiring them to do something they didn't think they could do. . . . I always recognize that anybody can succeed if they want to.*

During my observations of Anne's classes I was often struck by the banter that filled her room. She treats her students as mature adults and they respond to her with familiarity and respect. In her course outline, handed out to students on the first day of class, Anne writes,

> We're all nearly adults, treat others with the respect you expect them to show you.
>
> This year is all about making connections, forming relationships, and in doing so learning from each other. It's all about recognizing our strengths and weaknesses; it is about creating an environment, here in this room, and in this building, where we feel safe to step outside our comfort zones; take a risk and expand our horizons. There is no failure in this room. Even if what you set out to do doesn't work out, taking the risk alone means you have on some level, succeeded.

Anne, however, has been changing her teaching style in response to the pressure to ensure her school grades more closely align with her student's exam marks. Her focus now is much more intently aimed at preparing students for the exam than it had been when she first taught English 30-1. She comments:

> . . . . *What is my goal? As a teacher, is my goal for the kids to have fun, and think "English 30 was the best year, we had so much fun, it was great", or to say, "I was really well prepared for my exam. My teacher did her darndest to make sure that I wrote that exam and that I did well on the exam." I think it requires an essential shift in thinking where we go, "my responsibility to the students is to make sure that they do well on that exam, and also to their parents and to myself and the administration." A*

> *large part of my job is to make sure that those kids do well on that exam, and that is why I had to change. I mean, the principal the first year I was here said "Don't worry about the exam, don't even teach to the exam, just teach the class, make the kids enjoy it" and oh yeah sure, that really bit me in the butt, the next year. So I really changed, I had to. . . .*

Anne's critical reading of the exam has shaped her approach to preparing students for it. Consequently, she has moved from a more creative emphasis in her teaching to a more traditional focus on literary criticism and analytical responses to literature. Anne describes this transition:

> *We did less airy fairy stuff and more hard core (if there is such a thing with English). We did more reading, analyzing, writing, homework galore. I used to teach* Hamlet *really creatively and the students had to pretend they were a board for a small theatre, and they had to decide whether or not they could produce this play and they had to base it on* Hamlet *and how much was it going to cost, and could they cut scenes, and put this whole thing together, and it was a lot of fun. They had a wonderful time doing it. It was next to impossible to assess, and everybody did really well on it, and worked so hard on it. I am not sure how much they actually learned* Hamlet *that year. So now when we do* Hamlet*, at the end of each act they take four or five pages of notes. Now it says these are the things I want you to know, these are the things I want*

> *you to understand, this is the real meat and potatoes of this work, this is why it is important, this is what this means.*

Anne's comments suggest a tension within her approach to pedagogy. While the assignments she described fit well within a constructivist approach to teaching and a transactional approach to studying literature, it seems that Anne views this approach as being academic light. On the other hand, Anne seems to equate transmission approaches to teaching literature as being more academically substantial. This contrast is highlighted by her use of image: she defines her constructivist assignments as being airy-fairy and her transmission-based teaching as dealing with meat and potatoes.

Anne's teaching units reflect her new perspective. Each unit she provided to me for this research contained extensive research into expert interpretations of the literature being studied. Given this tension, when faced with the gaps between school scores and exam scores, Anne's response was to move away from a constructivist approach to the more traditional transmission approach that Anne associates with greater academic rigor. It is clear from the quotes above that Anne made this switch feeling that it would enable her to better prepare her students for the English 30-1 diploma exam.

In addition to being influenced by her diploma exam experiences, Anne's approach to teaching writing has been informed by her experiences as a journalist, by being a member of the diploma exam marking team, by her teacher training program, by former English language arts teachers, and by her personal writing style.

Anne completed her BEd as part of a two year intensive after-degree program. She feels that to some degree this intensive program did not serve her well. She felt unprepared for her first teaching job; at the same time, however, she felt very prepared in terms of her ability to teach writing. She reflects,

> *I do remember, though, that we did a lot of studying of how to teach writing, we did a lot of looking at different ways of teaching writing, we actually worked really closely with a junior high school, so that part I felt pretty confident about. . . . A huge component of what we did was writing process and teaching this and going into junior high and elementary schools and watching teachers teaching writing and evaluating . . . . [we] sort of beat the writing process to death—you know at a certain point you just say, "let it go," because the process is different for everybody, some of us plan more than others, and I think students as a whole don't plan at all. . . . I sort of figured that I knew how to teach writing. I didn't think I would be teaching English 30 right off the bat anyways so I wasn't that concerned about preparing anyone for a diploma exam. But we didn't talk about the diploma much, we talked about teaching writing and all the stuff that was involved with that.*

She left her BEd program understanding process-oriented approaches to teaching writing, she had direct experience working with students on their writing, and she began to understand the process students engage in when they write. In the years following her undergraduate training, Anne further developed her perspectives on teaching writing while working as a journalist. In this context she learned to emphasize the value of well developed ideas, the importance of brevity, and the effective use of diction.

> *I am always harping on them about the basic stuff, I have very high expectations, I expect them to produce a lot, not just volume-wise but also ability-wise, and I think they respond to that.*

Anne pushes students to expand and develop their thinking. One primary method through which she does this is by providing students with guidelines or formulas for structuring and developing their writing. She presents them with the formula, teaches them how to work with it, and then provides them with an opportunity to work with the formula in their own writing. Anne's *Hello Assignment*, handed out on the first day of class, reads as follows:

> In paragraph one, give me some personal information about yourself. You could include information like your full name, how your parents chose your name . . . . Tell me at least five things about yourself. (Worth five marks)
>
> In the second paragraph, tell me about you and school . . . . Tell me at least five things. (Worth five marks)
>
> In paragraph three, give me information about the activities and interests you have out of school. . . . Tell me at least five things. (Worth five marks)

In her units on essay writing, Anne focuses her students mostly on following the five paragraph essay structure. During my observations in Anne's classroom, Anne was focused on an end-of-unit essay assignment. She provided students with a five-page handout which explained the characteristics of the five-paragraph essay. Each page on the handout focused on one paragraph. For each of the three body paragraphs she writes:

> **Introductory Sentence**: (Note: May include transitional statement. Sentence introducing the point of this paragraph and the argument that will be presented.)
>
> **Supporting Sentences**: (These sentences will discuss your above stated topic ONLY. These sentences will use quotes to provide the reader with proof that your points are correct. The quotes must therefore be well chosen and clearly explained. This paragraph will end with a concluding sentence that sums up and relates this topic to the theme—do not forget to explain how this particular point relates to the theme.)

In her lecture to students, Anne elaborates on the formula. Her overhead notes on the conclusion, for example, stress the following:

> The conclusion . . .
>
> Upside down introduction
>
> 1. restate you thesis
> 2. remind the reader of your essay's focus
> 3. does not begin with "in conclusion"
>
> Reminds the reader of the most effective details
>
> Contains universal appeal—shows how your ideas have purpose outside the literature—the big picture—
>
> Does not contain any new information

Springboards from the technique you used to open your essay.

This approach to focusing students on formulaic structure is commonly used to help them develop a sense of organization and idea development. Some criticism of this approach is that it focuses the writer too much on the structure and not enough on the ideas. Anne acknowledges this concern through her constant reminder to her students to "dig deep" while developing their ideas.

In addition to assisting students by providing them with formulas upon which to base their writing, Anne provides them with direct support and coaching. She is actively involved in assisting students with their editing.

> *In my case the editing process is brutal. For those kids who are weak, they don't even realize that what they've written is substandard. They don't even have a clue. So I really feel that it's my job to sort of be the nasty editor and say, "ok, keep these ideas, these are great but all the drivel around it, why did you bother with this?"*

Anne works with students on editing and revising skills. She asks students to engage critically with the writings of other students (classmates or anonymous exam writers) and with their own writing. She hopes that these editing skills will help her students independently improve their own writing. Though she feels that too often students seem to lack interest in this element of process. Her experience with her students is that they merely want to get their writing done as quickly as possible.

Anne's approach to teaching is also to some degree influenced by her own writing process. Anne believes in the value of planning and organizing prior to drafting. She is concerned that her students do not seem to engage in this process, though she designs her assignments to provide time for process. She comments,

> *When I give [my students] an essay I give them a long time to do it. I give them a week to do it because I expect them to do a couple hours here, an hour and a bit there, and another hour here, leave it for a day, come back to it. I find they don't spend enough time planning, they don't spend enough time organizing their ideas, and then a lot of them don't hand in first draft writing, but they definitely don't have the editing skills to take it from rough to polished.*

Anne gives her students time in class to work on their writing assignments. She focuses this class time on supporting and developing writing process. Anne, however, recognizes that in her own writing she does not follow the same step by step process that she requires of her students. She attributes this difference between what she does in her own writing and what she expects her students to do to the diploma exam which requires students to generate and organize ideas quickly prior to writing.

> *My own process is probably more haphazard than I teach. I think I'm more of a Type A so I don't necessarily outline, but everything has to be properly paragraphed and 'T's crossed and 'I's dotted and all those kinds of things. . . . I take a long time to organize the paragraphs and make sure that they can follow the thought process that is happening there. When I give them assignments I do have them outline only because I think it's one of the ways that they can put their ideas together quite simply and quickly*

> *without having to worry about the grammar or the sentence itself, but just getting the ideas down quickly because that's something that they're going to have to do especially for the exams.*

Anne is flexible in her approach to teaching writing process. Each genre, she suggests, requires a different process and individual students benefit differently from a range of processes. Anne's approach is designed to assist students in developing an important skill needed to perform effectively on the diploma exam while also enabling her to provide opportunities for students to engage in their own, unique approach to process.

Anne sets time aside for students to work in class on their writing. She spends several periods at the beginning of the assignment talking about key ideas that students can work with in their writing. During this time she also provides an opportunity for students to write in class. Anne then expects students to complete their work at home and have it handed in several days later. She provides class time in the early stages of writing process. This reflects her emphasis on the value of preplanning and idea generating prior to drafting. Time spent in class at the end of the process would enable Anne to also work more with her students on revision and polishing elements of writing process.

In spite of the freedom to work out their own process and the time provided for students to engage in a process-oriented approach to writing, Anne feels that her students most often engage in a simplistic writing process. She reflects,

> *My own experience with [my senior high students] was that they just want to get this thing out as quickly and painlessly as possible, they don't really care about spending the time on it. They're not looking at writing a great*

> *piece. And I think part of that is, they are taught the proper process. I think in the junior high grades, at least in our school, I know the language arts teachers go through every single step, I mean the writing process becomes a belabored issue with the kids and I think that it is partly what turns them off, they think "I can skip all that other junk and just get to the writing bit." So I think it is almost as much an attitude. . . so they are getting taught and then I think they make a choice, "Oh I want to get this done, I am going to pick and chose what I do with it, and go from there"*

Anne suggests that students' choices regarding writing process reflect the degree of commitment they have to the assignment, the grade they wish to receive, and their previous experience with writing process. Anne suggests that while frequently presenting students lengthy discussions or lessons on writing process provides students with important information on how writing process functions, she also suggests that this approach turns students off process, compelling them to focus merely on getting a draft done as quickly and painlessly as possible.

Anne's approach to teaching writing is also influenced by her experience marking diploma exams for the Alberta government. Marking diploma exams for the first time—several years after she had decided to modify her teaching approach to better prepare students for the exam—was a positive experience for Anne. She comments,

> *[Marking diploma exams] was great. . . . the experience itself, [gave] me the confidence to know I was on the right track, to know that my kids could*

> *produce, and that what I was teaching them was actually valid and working and was going to be appropriately rewarded on the diplomas. . . .*

Anne's experience marking exams helped her gain insight into how her students were performing in relation to students across the province, it helped her understand what skills or knowledge were important for performing effectively on the diploma, and it helped her understand what knowledge and skills were unimportant. This bank of knowledge helped Anne regain some of her confidence. Her previous experience with the diploma exam had been troubling for her, causing her to question her competence as a teacher. She reflects:

> *The test is something that is given province wide so we feel, well, if we taught our kids what they need to know then they will do well enough on the test, they'll be close to what we assigned in the classroom. I think we kind of see them as a performance indicator, as a measure—I certainly felt really strange when my exam marks came in and they were twenty percent lower than my classroom marks, that was quite alarming. I sort of thought, "Holy, does that mean I am a total pushover as a teacher, that I am assessing the wrong things, or that I am not marking fairly, or properly." I mean I had all these doubts about my abilities as a teacher based on those exam scores.*

Notions that exam scores are a reflection of a teacher's ability are reinforced in publications such as the *Fraser Institute's Annual School Rankings*. In their 2004 report, for example, they cite Hoxby (2002) who claims that, "statewide standardized tests and school report cards may be unpleasant for ineffectual educators, but they should not be controversial with parents or policymakers who want to see higher achievement (p 2)." Hoxby's comment constrains the discourse around test-based accountability. Essentially she implies that those who have a problem with test-based accountability must be ineffective educators; effective educators have nothing to worry about because the tests will demonstrate their competence. While I disagree with Hoxby's assertion, I do recognize that her perspective is shared by many researchers, teachers, and members of the public. Anne seems conflicted about her perspective: on the one hand her statement above suggests that she sees her students' exam performance as a measure of her teaching ability, but on the other hand she recognizes that contextual factors have a significant impact on her students' performance. These factors make it difficult to attribute exam scores to teacher competence. In dialogue with Heather (a third participant in this research) Anne discusses her frustrations:

> ***Anne:*** *That's how my year always starts. The first year with this principal [he asked], "Why was there such a discrepancy between your class and the diploma marks?"*
>
> ***Heather:*** *But you can show them why.*
>
> ***Anne:*** *Absolutely, because we did speeches, because we painted pictures, because we wrote, because we talked, because....*
>
> ***Heather:*** *So he has his answer, then the question is done right?*
>
> ***Anne:*** *Well for the most part. But the expectation, the implication though is that they need to be more in alignment. And then I've had parents come*

> *back after the diploma exam – why was there such a discrepancy between my child's mark and their school mark? Parents come and seek me out.*
>
> **Heather:** *Did your kids have exam anxiety? Did your kid get enough sleep that night?*
>
> **Anne:** *There were a number of different things. I mean this kid failed every single diploma that she wrote. She had 80 in every single subject and failed every single diploma. So what else is going on here, why is it my fault? But the implication is that there's something wrong.*

Anne's discussion with Heather reveals Anne's frustrations: she is being judged on her ability to teach English language arts on the basis of an exam which purports to measure students' writing ability, but she has no control over how well students perform on the day in question. While Anne recognizes her professional obligations as a teacher, she also expects that students take some responsibility for their learning. She suggests that socio-economic and cultural backgrounds of student populations have an impact on average school performance. She resents being made to feel responsible entirely for student performance even though there are significant factors relating to student performance that she cannot control.

Anne attempts to base her planning half on what is expected in the program of studies and half on what the diploma exam is designed to measure. Anne's units are mostly focused around themes. The first unit of her English 30-1 course focuses on isolation as a theme while her second unit focuses on civil disobedience. Each unit consists of up to ten literature selections which range from song lyrics to poetry, to short-stories, to feature films. Student writing in Anne's course is almost exclusively structured as creative or critical responses to literature. Anne's unit on isolation has students complete journal entries, plot outlines, two personal responses to literature (modeled after diploma exam questions), short answer questions focusing on close reading, a character profile assignment, a graphic essay, vocabulary sheets, and a story board. Clearly Anne is quite diverse in her expectations of her students' writing. She requires them to develop skills related to creating a broad range of genres. In spite of this range, however, her focus remains fixed on the diploma exam. Her units' themes reflect previous diploma exam questions; and the final assignment for each unit, a critical response to literature essay, is modeled after the diploma exam. Reflecting on her writing program, Anne suggests that all her writing assignments are geared toward the final exam. She elaborates:

> *I would say that every single writing assignment I give is in some way directed toward the final exam, even a journal entry, because a journal entry allows them to create their own context and personalize what they are reading which is what they need to do in the personal response. . . . Does it influence the methodology? When I am teaching critical essay writing, absolutely. That critical essay makes or breaks my students. Most of them will do fairly well on the personal response (very well being 65% or better) but they will bomb the writing test even just from fear alone, they'll just seize up and that will be the end of it.*

In addition to designing writing assignments to focus on the type of thinking and writing that students need to engage in while writing the exam, Anne also has her students focus explicitly on the exam itself and on the exemplars of student exam writing that Alberta Education provides to students and teachers. (These exemplars are designed to illustrate the exam's standards.) Anne has her students deconstruct previous exam questions, deconstruct each other's essays using the diploma exam marking guides, analyze sample student responses to previous exam questions, and within the context of this work she engages them in conversations regarding the subjective nature of assessing writing. She describes her approach:

> *[The exam] directly influences the way I teach writing. You know when we look at the essays we do a lot of dissecting of other people's essays. So I will have them writing essays and they will have to go around and use the rubric, the actual [exam rubric] that they give you when they mark the test and then they have to score each other's papers. I think that is the best way for them to get a feel for what is good writing and what is not good writing. We do a lot with the examples that are on the Alberta Education website. We take them apart, "what makes this good, what doesn't make this good?"*

This approach, she believes, helps students to develop a critical awareness of their own and their peers' writing. This critical engagement should help them to develop an awareness of the weaknesses in their own writing, and it should encourage them improve their writing.

In addition to designing writing assignments and other instructional activities to directly focus on diploma exam preparation, Anne's assessment of student work is also influenced by the diploma exam. She uses the diploma exam marking guide when marking student writing. She encourages her students to use this marking guide when evaluating their own and their classmate's writing. Anne uses the rubric both because she feels it is well designed and because she feels that consistent use of the rubric will help her students understand what kind of marks to expect on the diploma exam. Anne values the diploma exam marking guide because it provides guidance in terms of channeling a marker's subjectivity. She comments,

> *I think the rubric itself is very fair, I spent five days with the darn thing marking exams with it in January and I actually found that it was very fair. Assessing writing is so subjective and even within the delineations that they have given you in the rubric, there are a lot of subjective—what is the difference between thoughtful and illuminating? There is a fine line right—so I thought that it was fair if you can get a feel for what those indicators mean. And that is why marking exams in January was critical for me because it really allowed me to predict "Ok so this is illuminating, this is thoughtful* [keywords in the exam rubric used to differentiate between levels of writing quality]. " *I think it is very fair, I think it is very accurate.*

Anne's experience using the marking guide while marking provincial exams helped her to understand the standards associated with the rubric descriptors. This experience has helped Anne to become clearer in her expectations for her students' writing; it has helped her to become more rigorous in her own marking.

Perspectives on the Diploma Exam

One might think, given Anne's focus on preparing students for the diploma exam, that Anne would be positive about the exam itself. Yet this is not the case. Anne is quite critical of the exam. She recognizes the limitations in what the exam is actually measuring.

> *The exam only tests a very limited writing style, structure. I mean you've got the personal response, you've got the critical analysis. For all the other things we do during the year, these things aren't even remotely tested at all. . . And that's ok, I mean yeah, there isn't time or ability to test them all. So I mean they've picked these two forms so I think that's where as a teacher my emphasis would go because I knew that was what was going to be. . . . If they're going to spend the most time on trigonometry on a math exam, that's where you spend the most time in the class.*

Based on this analysis of the exam's limitations, Anne is able to more directly focus her instruction on developing student competence in relation to what the exam is explicitly measuring. Anne does acknowledge, though, that there are many forms of writing that her students are asked to do in class that are not measured by the exam.

Anne is conflicted in her perspective on the diploma exam. On the one hand she values the idea of standardized diploma exams because of their perceived ability to create a level playing field for students applying to university. While acknowledging this benefit, she suggests that the exam suffers from serious construct flaws. She comments,

> *I don't necessarily feel that [the English 30-1 diploma exam] is a fair assessment, I think the writing component of the diploma itself is probably the least indicative of what a student can do, it is pressure writing, it is writing out of context, I mean for all the things we teach writing to be, it is not, it is the opposite of everything we want it to be. . . . Is it a measure of what a student is capable of? Yes and no. The strong students who do well under pressure sure it is. But you know, the majority of us go through life in the mid range and we suffer from a fair degree of test anxiety and I think that that is a real factor for a lot of kids so in that sense I don't think it is a fair measure of what a kid is able to do.*

Anne criticizes the exam because it operates within a high pressure, high stress artificial environment and as such is measuring student ability to handle stress. She also argues that the exam is not providing students with sufficient time to engage in even a limited writing process. And she thinks it is unfair to students that the difficulty of the exam changes as the exam question changes. She comments:

> *I don't think [the diploma exam] is accurately measuring how well the kids can write. It is measuring how well can kids write under extreme pressure, because they are not comfortable with this [the June exam] topic, there is no where they feel like they can go with this topic. . . . They aren't relaxed, they can't come up with anything, they're trying very*

> *hard—focus, organization, these kinds of things are going to be totally out the window as far as I am concerned. . . . And I did have students say, "I didn't finish, I didn't have time to finish," because they had to spend so much figuring out what to write about. So I feel that it was a very unfair assessment. . . . [The exam is measuring] the fear factor, obviously. They're panicking, they're reading the first selection and they're starting to panic.*

In spite of these criticisms, however, Anne is quite clear that she uses the exam to guide her planning, her assignments, and her assessment practices. She recognizes that this seems to be contradictory but explains that given her context she has little choice:

> *It is a complete paradox, you can't do that right, you can't say the exam is not a fair assessment but I am going to use it anyway. But that is what we do. . . . Because it is really hard to reconcile those two things. You say well, the expectations of the administration, and of the parents, and of the students themselves, is that you prepare me for the exam. Okay fine, I can do that but personally I don't feel that this exam is a fair assessment. But that is the expectation, so then I have to balance these two, and wrestle with these two in the classroom, and say okay, "I am going to let the exam go a little bit and we are going to do something wild and creative and have a little fun with this piece of literature rather than focus exclusively on the exam." It is really hard to reconcile those things together. . . . Yeah, valid or not, the darn thing still exists, and students are going to*

> *have to write it. I mean the only way you could eliminate the tension is to eliminate the exam, and that is not going to happen, that is not going to happen.*

Anne feels that she has little choice but to teach to the exam; her administrators, her students, and her school community at large expect that this will happen. Anne has learned that her administrators and her school community judge the quality of teachers based on student performance on diploma exams. She had trouble with this her first year teaching English 30-1 and has decided that she does not want to have the same experience again.

Brian's Profile

Brian is in the eighth year of his teaching career, a career he began in his thirties after completing his Bachelor of Education Degree with a major in Math and a minor in English. He teaches English language arts, mathematics, and science. Brian teaches in a K-12 independent school on the edge of a major city. The school draws its students from both the urban center and from the acreages surrounding it. Students at this school generally come from middle to upper-middle class backgrounds. The student population at this school is quite stable, averaging between 180 to 200 students. Brian teaches courses ranging from English 30-1 to Math 9, to Science 5/6; and so, by the time he teaches students in English 30-1 he has known them for a number of years already. He considers this to be an asset in his teaching for it provides an opportunity for him to forge good relationships with his students.

Brian's classroom varies in terms of its organizational structure. With 16 students in his English 30-1 class, the room is quite spacious. Two large bulletin-boards cover the rear and side wall of the classroom. To varying degrees, these are filled with student work, mostly posters. In the back corner, opposite Brian's desk, is a framed water-color painting produced by a former student of Brian's. The poster represents the student artist's reflection on *Macbeth.* While desks are arranged in rows facing the front board, students frequently move their desks to support group work or to assist one another in their work. Brian's desk is in the back corner of the room. During class, students will occasionally stand or sit next to him as he discusses their work with them. At other times Brian circulates through the room stopping to assist students who request his help. Brian suggests his room is marked by what he calls "productive noise," the low hum of energy one hears when students are working on projects and assignments.

Brian's alternate classroom is the school's computer lab. Approximately half of Brian's classes that I observed took place in the computer lab. Brian provides ample opportunity for his students to work in class on their assignments and projects. Having students work in the computer lab enables them to conduct internet research and to engage in word processing activities. The computer lab is spacious. It has 24 computers lining three of the room's perimeter walls. At the center of the room are three rows of tables.

Brian is committed to ongoing professional growth and development. He is a member of the Alberta Teachers' Association's English Language Arts Council (ELAC), and regularly attends its annual professional development conferences. Many of the assignments he utilizes in his English 30-1 program are drawn from sessions he has attended at ELAC conferences. As an English minor, Brian's undergraduate course-work focused more on mathematics than it did on English language arts. The English courses Brian did take focused on various aspect of studying literature. He claims to have received limited explicit instruction in composition or in methods for teaching writing. Consequently. Brian feels that he really struggles in his teaching of writing. He comments:

> *Teaching writing is [something] that I really struggle with a lot because, how do you teach writing effectively? I know for myself. . . I don't feel that I have a real great way of teaching it. I do know that I can identify things that the kids can improve on and I try to work that into writing units but when it comes to [teaching] writing I know that there is a lot of room for improvement. It is really difficult teaching writing. . .*

Brian, who coaches junior high basketball, frequently uses sports metaphors to describe his frustrations and challenges with teaching writing. The metaphors are apt because they focus attention on the fact that writing is a skill which needs to be developed rather than a base of knowledge that needs to be transmitted. Brian comments:

> *If you are playing basketball the only way you get better is by practicing and doing it, you are not going to get better sitting in a classroom having someone tell you, "okay this is how you take the ball, this is how you dribble it". You have to be on the court doing it and that is the same with*

> *writing. That is one thing I have learned over the years—I am an English minor, I don't have a strong background in English, I haven't taken a lot of writing courses—I know that to have a grammar lesson—"today we are going to cover the comma"—that in itself doesn't make them learn how to use it. They have to learn how to use it and do it.*

Brian's coaching metaphor influences his approach to pedagogy as well. He sees himself as a coach, providing students with the time and opportunity to write while also being there to guide students into developing into more effective writers.

> *In my writing units, what I have done—well last year for a while, we haven't done it this year—we had weekly journal entries and then a formal three week writing block where I just assign four or five essays that have to be written within that time period. The class time was almost strictly used for writing in the computer room—work on your writing—I am available to conference with them if they have problems, even if they don't think they have problems, I know who the weaker writers are so I will pull them aside and say, "Let's see what you have written so far. . . ."*

The assignments Brian chooses, the structure of Brian's classroom, and the structure of Brian's lessons collectively support this coaching approach to teaching writing. While Brian does suggest that time constraints significantly impede the degree to which he might follow through on such strategies, specifically the number of writing assignments he gives to students, he certainly focuses his instruction on providing students with opportunities to practice their writing in a range of contexts.

Descriptions of Practice

Though Brian is confident in his knowledge of the English language arts curriculum, he is constantly challenging himself to find new and interesting ways to engage his students. To this end, he frequently varies his pedagogical approach, shifting between lecture, in-class writing assignments, and project-based approaches. Particularly, Brian enjoys project-based teaching because he feels that it engages students.

> *But I also like finding projects that the kids really enjoy doing too. I'm fairly big on projects, a lot of visual representation and things like that. . . . I like finding projects that get them into a topic and then you can work with them. Then they don't mind a bit of lecture and they don't mind doing some writing.*

Brian's major writing assignments, administered throughout the year, also reflect a range of teaching styles moving between student-centered and directed to teacher-directed. One stream of assignments is built around the theme of "Writing Hints." These assignments focus students on explicit tools for writing. For example, his assignment on transitions has students focus on four ways to make transitions within their writing, and his assignment on sentence structure has students focus on coordinating conjunctions, subordinating conjunctions, compound sentences, complex sentences, and compound complex sentences. These assignments are designed using a worksheet format. They require students to create examples of proper usage but do not require students to reflect on usage in previously developed work.

To a significant degree, Brian's writing assignments are embedded within a response to literature focus. This design reflects both the structure of the *Senior High ELA Program of Studies*_ which emphasizes the inter-relation between the various language arts, and it reflects the design of the English 30-1 diploma exam. Brian is very aware of the responsibility he faces for selecting literature that offers students a range of topics and choices for response:

> *You choose your literature to a degree so that it will work for the exam . . . and if you are not sure you go in picking literature that might not help the kid at all for an exam question and that would be unfortunate since [the exam] is weighed so much.*

He is acutely aware of the consequences for his students if he chooses literature for his course that doesn't offer the possibility for students to engage with the text in relation to a range of questions. He is also aware that students need practice and experience with thinking about writing about literature in the manner framed by the exam:

> *I will also give some assignments throughout the year that are very closely related to the departmental exam questions—the personal response, but also the more structured, longer response—so they will be prepared for those types of questions when they actually get to the departmental exam. . . . I don't do a lot of real exam prep and I probably should do more just to*

> *get them comfortable. I do take exam questions from the past and I'll get them to write an essay: there was one personal response from January 2000, I gave them that as a personal response they had to do in class; I have also taken questions and made them more specific by applying it to an actual piece of literature. So for* The Great Gatsby *I had them write a unit ending essay and I took an actual diploma exam question and rather than leaving it generalized I specified they had to answer with* The Great Gatsby*. So that way they are at least comfortable with those types of questions.*

Brian's comments demonstrate that he is not ultimately interested in teaching to the exam, but it is clear he feels a professional obligation to ensure that his students are comfortable with the exam and are prepared for the pressures and demands that the exam places on them.

Within Brian's literature-based assignments are main groupings of similar assignments. One group of assignments are entirely teacher directed. For example, an assignment designed as a response to *A Tale of Two Cities* asks students to:

> Read the opening chapters to Hard Times by Charles Dickens. What aspects of the Dickensian style can you find in them? Using two chapters of your choice from A Tale of Two Cities, write an essay on the *Dickensian Style* . . . In writing your essay, you can impress me by finding an interesting way to address the topic.

This assignment like several others administered throughout the year is clearly focused on a teacher-chosen topic. The other teacher-directed assignments, however, are more focused on student interpretations of text, rather than being focused on elements of style. An 80 minute essay test on *All Quiet on the Western Front*, for example, asks students to do the following:

> Address the following topic in a well-written essay. Be sure to focus your essay, and to organize your thoughts in a logical manner.
>
> > Write an essay based on All Quiet on the Western Front, in which the author creates (or recreates) situations, characters and an event which are meant to color the reader's perspective. What angle does the author develop in this novel and how does he support it? Provide specific details from the novel that will support and develop your controlling idea.
>
> Use the paper provided. You may use your text to help you with the essay, but you may not use any notes, etc. You have 80 minutes to complete the essay.

While the question itself is quite specific, it does provide students with an opportunity to engage with the text from a number of perspectives and angles. Conceivably this assignment could generate a range of essays which each focus on a different angle and/or set of events or characters. Other similarly designed literature-based assignments differ from those presented above only in terms of student choice. For example, an essay assignment on *Hamlet* provides students with a choice of one topic from among nine choices. Choice nine being, "if you have another idea for an essay based on *Hamlet*, let me know."

The second set of literature-based assignments is more student-directed and project-oriented. For example, an assignment designed as a response to the short story "Paul's Case" requires that students first answer a series of 26 questions which each ask students to think about Paul in terms of a range of metaphors (eg. "If your character was a color, what color would your character be? Why?"). Students, working in groups, are then asked to pick the 12 strongest metaphors from their list and develop a collage, diorama, poster, or any other type of visual representation which represents their perspective on the character. Students are expected to develop a meta-cognitive reflection essay which explains the image and the choices that went into creating it. Similar assignments include the design of a new cover and a new title for *A Tale of Two Cities* (including an inside jacket which contains a biography of the author, and a back cover, which includes fictional endorsements and a brief synopsis of the novel). Projects for *Macbeth* include several choices: creating a poster advertising a new movie based on the play; a collage related to the play's themes, or a collage which draws attention to common phrases used today which find their origins in Shakespeare's works. Each of these collage assignments requires students to complete a meta-cognitive reflection essay explaining the thinking behind their creation. Time lines for these types of assignments range from four days to two weeks.

A final set of writing assignments are not framed around literature responses but rather are focused on various genres or aspects of writing. The most comprehensive of this set of assignments is a five week essay writing unit. During this unit students are expected to write four essays, each due a week apart. Essays are expected to be 600 to 750 words long. Brian describes the assignment to students as follows:

> Your essays will be topics of your choice. We could brainstorm some topics together as a class if you find that will be helpful for you to find a topic. All English class periods in this five-week period will be given to you for writing your essays. I will be available to answer questions you might have, and you may also get advice from your classmates. You will be expected to come to class prepared to work. We will have access to the computer room for most of the classes.

The assignment offers Brian extensive opportunities to assist students in their writing because they are completing their writing in class. Though, my observations, on a similar unit suggest that students spend much of their time on the preliminary elements of their writing process, but do not spend much in class time talking with Brian about revision elements of process.

Though many of Brian's assignments incorporate creative requirements into them, few of his assignments call for students to engage in creating the types of genres that typically are considered to be creative writing: short stories, poetry, scriptwriting… Brian attributes his selection of writing assignments to his own writing style and comfort levels. He chooses essay assignments and more formal types of writing for his students because he sees himself as a structured rather than creative writer.

> *It is very easy for me to direct them to the style that I am comfortable with but that isn't what is best for them always. . . . I have a student in grade 11 right now, who when she wrote her achievement test in grade 9 addressed the question in a short story and it was just phenomenal and I knew it would be because she is just an incredible writer. When she tries to write structured . . . she scored very average marks.. . . She doesn't fit that very structured kind of approach and that is where too I have to remember that when I teach writing also, she's got a different style and it is a wonderful style, she is a phenomenal writer. . . . She is such a great writer and I'd hate to put her into that structured box of mine.*

While Brian acknowledges an emphasis on structured writing in his English 30-1 course, it is interesting that his assignments allow for creative freedom. He does not insist on specific structures or formats for his students' writing, rather, his assignments leave possibilities open for his students to explore structures that are appropriate to their own ideas. The same can be said for process; Brian's assignments allow for students to engage in their own form of process rather than rigidly adhering to an artificial conception of process. In fact throughout the interviews, Brian suggests a relationship between form and process.

> *I find that when you teach process, and the most common process in junior high is introduction, three paragraphs and a conclusion, and that works for that stage in their learning but I find . . . it is very difficult to get them away from that once they are in grade 11 and 12.*

This conflation recognizes that predetermined structures often lead to predetermined step-by-step processes. If you are to write a five paragraph essay, you begin with the introduction, move to defining your topic sentences and key ideas, write your body paragraphs and so forth. While Brian does not emphasize this approach to writing in his own teaching, preferring instead to allow for more organic processes and structures, he does suggest that there are benefits to teaching children to use a rigidly structured approach to writing:

> *If that is the way they are most comfortable writing well maybe that is fine but what about those students who aren't like that . . . , I don't think many of them get beyond that basic structure. . . . I don't know if I would change the way it is done because I think that basic [5 paragraph] form is a good form, because it teaches organization. . . . So I don't think we can go away from it, but maybe when we are teaching it we can maybe tell them why we are doing it that way, tell that it is not the only way.*

During my four weeks of observing Brian's English 30-1 class, Brian's students were completing a poetry portfolio assignment. For this assignment students were expected to do the following:

- Collect 9 poems written by Canadian poets.
- Research and write short biographies on 3 poets whose poetry was represented in the 9 poems selected.
- Write personal responses to three of the poems
- Write technical responses to three of the poems.

- Write a critical response comparing three of the poems.
- Present each of the poems artistically – using background image and font choices to reflect the content or intent of the poem.
- Create several original poems of their own.

The assignment is comprehensive, covering most of Brian's work with poetry during the course. The assignment is designed around Rosenblatt's (1938) transactional theories of reading, asking students to engage with texts from their own perspectives, first personally, then technically, then critically. Students are given significant choice in determining what poems they wish to work with and what responses they wish to generate for each poem.

Students spent most of their class time working on initial research. One third of the students took the whole first week to find their poems and to decide which poets they wished to write biographies on, several students were working on this material in the second week of the assignment. Many students also spent a considerable amount of time completing the creative elements of their portfolio, specifically the poem presentation pages. Creatively presenting the poems they had chosen also took considerable amount of time for the students. While Brian encouraged them to complete some of this work at home, and to work on drafting their personal, technical, and critical responses in class where he could help them, a large number of students decided to work on the creative presentation of poetry during class-time. This choice limited the amount of feedback students could receive on their actual writing during class.

Brian implemented the assignment with a number of goals in mind. During our first interview he described these goals as follows:

> *One of my goals for the poetry portfolios was to get the kids to read poetry and then have [them] personally respond to the poetry—that, I guess, could be tied directly to the diploma exam in the sense that it always used to be [a] poem [they had to respond to]—then we do the technical response which makes them analyze poetry. So those are the three: to read poetry, to personally respond, to technically respond. They do a lot of writing. They [also] write a couple of poems. And then the creative aspect of it. . . : I want them to present [their poems] by working with the theme, or a message, or some strong imagery. So in all aspects [this assignment] makes them work with the poem, it is a pretty in depth assignment. . . . I found with a lot of the kids—not everyone, but a lot of them—is that the finished product they are very proud of because they have put a lot of work into it. I know kids that I have taught, that have done this four years ago, who still have their poetry portfolios because it is something they are proud of.*

The final product students create for this assignment is often a 20-30 page document. The assignment aligns well with the diploma exam's expectations for student skill development. The diploma exam focuses firmly on written responses to literature. As such this assignment, while not directly an exam preparation activity, is designed to assist students in developing a set of skills being measured by the exam.

Brian's end of unit tests and his mid-term exams strongly reflect the diploma exam design. Half of these tests or exams require students to answer multiple-choice questions drawn from previous diploma exams. The second half of these test/exams require students to write a personal response either to a text studied in class, or to texts provided on the exam/test.

During the interviews Brian suggested that the exams did not overly influence his teaching of writing. He explains:

> *When planning my English 30-1 course] I think I use both [the Diploma Exam and the* Senior High English Language Arts Program of Studies] *because....the* Guide to Implementation *is actually a pretty good guide. . . : I did a heroes unit this year, the idea came out of the* Guide to Implementation; *I didn't' use a lot of the things out of it but the idea came from there. But then I also used the diploma exam model. I'll grab a personal response question once in a while and fashion it to a novel that we're studying, or work that we're studying, so that they get. . . an idea of what kind of question they're going to get.*

Brian's writing assignments certainly reflect this approach. His assignments are designed to address skills or content expected within the *Senior High ELA Program of Studies*. For the most part, these assignments do not reflect the diploma exam design too closely, though many of them certainly help students develop the response to literature skills required by the exam. These skills are also important skills the program of studies expects students to develop in English 30-1. A small number of Brian's writing assignments are closely modeled after the diploma exam questions and format. Brian explains that the primary goal of these assignments is to help students become familiar with the format and questions they can expect when writing the diploma exam. A majority of Brian's writing assignments, however, focus more broadly on curriculum goals and expectations.

The marking guides accompanying these writing assignments are also varied in terms of their design, focus, and degree of detail. For example, *A Tale of Two Cities* book cover assignment has the following marking guide:

A Tale of Two Cities Book Cover Assignment
Name:____________________

Front Cover:	/10
Back Cover	/10
Inside Cover	/20
Spine	/2
Total	/42

Other than providing the weightings for various components of the assignment, this marking guide provides little indication of specific marking criteria. On the other hand, the marking guide for a "Research Paper Assignment" lists four major categories—research, management of information, communication of information, and conventions of language—that each contain three levels of descriptors and up to five specific criteria.

The focus of this rubric is technical in nature, drawing students' and the marker's attention to issues such as citations, diction, and organization. Brian's standard rubric however, is modeled after the former English 30 diploma exam rubric. Brian's decision to use the diploma exam rubric for his own marking is quite deliberate. He explains:

> *[Marking my students' work] is probably where the exam comes most into play, because basically I set up my rubric along the same lines as the departmental. . . . It is helpful for me because it helps me to focus on what I should look for, but it is good for the students so that when they get to the exam they are not going to be facing a whole different marking style. . . . [When I see the exam results] I am glad when I am not way off (or that the exam is not way off) from what I perceive to be where the kid is at, . . . that is partly why I used the [diploma exam]: . . . I want to mark the kid where they're going to mark the kid.*

In addition to the benefits for the teacher, Brian's use of the exam rubric enables his students, over the course of the year, to shape their writing to reflect the focus of the exam rubric which could in turn lead to better performance on the exam itself.

Perspective on the Diploma Exam

Positive attributes. Brian's perspectives on the diploma exam are complex. As discussed earlier, Brian does not allow the exam to explicitly drive his teaching, yet at the same time he feels a professional obligation to his students to prepare them for this exam. Primarily, Brian values the diploma exams because he feels that they create level playing fields for students across the province who are applying for post-secondary education; they hold teachers accountable to the curriculum; they validate teaching; and they provide an objective indicator of student ability. Yet Brian is not fully supportive of the English 30-1 diploma exam program. He feels that it contains significant flaws: it has construct problems and its stakes are high making it tempting for teachers to narrow their curriculum to that covered by the test. In expressing this balanced view, Brian says,

> *I've never had a problem with standardized tests. I think they can be a problem if a teacher makes it absolutely rule the way they teach. But I do agree that there is room for standardized tests because there has to be some guard against teachers going their own way and really just teaching whatever they want to teach and really not giving the kids what they need. I think that you can work with them. . . . In the end I do appreciate the diploma exams in the sense that I think it keeps teachers honest and it creates more of a level playing field for students who are going from so many different schools and who are applying for positions at the same universities; I think definitely there is a place for them.*

Brian's perspective is informed by his own high school experiences. He too went to school in Alberta but during a time when diploma exams had been phased out. He recalls, in particular, a high-school social studies teacher who had an interest in communist ideology. He recalls that she built her entire course around the topic with little regard for the actual curriculum expectations. Brian is certain that had a diploma exam program been in place, this hijacking of the curriculum would not have occurred.

Brian is also concerned with grade inflation. He knows that teachers feel the pressure to ensure that their students receive good grades. He knows that in Alberta there is a certain pressure on teachers to ensure that their school marks and the exam scores are within a few percent of one another. This is supposed to guard against grade inflation. Brian comments,

> *The principal likes it when the test score comes in close [to the school marks], and I know from talking to other teachers that I don't get the pressure that other teachers do. There are some teachers who are under a lot of stress to make sure that their class marks come in [within range of exam scores]. They are under a lot of pressure to have those marks come in very close. I don't have that here, I know our principal likes it, but I mean if I was five percent off or seven percent off I am not afraid of losing my job.*

In part, this pressure to align grades is external, stemming from the way Alberta Education reports diploma exam results. Their results tables indicate for each school, the average school awarded mark in one column with the average diploma exam mark in a second column. School results are also published in local papers usually within the context of the Fraser Institute's report cards on schools. And schools themselves often report performance on diploma exam and achievement test results. Though Brian does not feel much external pressure from either his administration, school board, or school community, he does place some pressure on himself to ensure that the grades are in some relative degree of alignment.

> *I think that generally [the spread between diploma exam and school awarded marks] should be close, I mean some kids [marks] won't be because they can't perform under that pressure and some kids can soar under that pressure. . . . If I am ten to fifteen percent out then I say, "ok I have to be a little careful with that". You want to make sure you are not overly inflating someone's grade in the sense that you give them some false sense of competency or ability because from here they do face the harsh realities of post-secondary education.*

Brian has struggled with criticism about his marking from some of his students who have accused him of being too subjective, of marking a paper based on who wrote it rather than how well it was written. For this reason, correlation between school scores and exam scores is important to Brian: a high degree of correlation enables him to fend off this criticism by pointing out that his school marks are reflected in the diploma marks.

Brian also uses exam scores for diagnostic purposes, comparing results across exams to determine areas where improvement is needed. He observes,

> *Our kids do score very well on the reading comprehension [portions of the diploma exams and achievement tests] but poorer on the writing—still above the provincial average but poorer than the reading. So we know there is a deficiency there. So we have gotten together as staff, and we've said, "we have to get these kids writing, we have to get them writing in*

> *their journal every week." . . . I know that our deficiency is in writing and I am consciously putting an effort on addressing that, and that is because we see the reflection in the test scores.*

Based on this analysis of results, Brian's school is dedicating its new round of Alberta Initiative for School Improvement (AISI) funding toward developing a systematic approach to improving writing instruction in the school. Some aspects of the project include team-teaching select English language arts classes at the elementary and secondary levels, completion of a graduate course on teaching writing by one staff member, and the granting of release time for another teacher (Brian) to research and coordinate improvements to writing instruction. This is an important initiative for the school, one that recognizes the value of analyzing exam results.

Critical perspectives. Yet, in spite of the benefits offered by the diploma exam program, Brian recognizes flaws in the English 30-1 diploma exam. Primarily these flaws relate to construct issues: The exam measures only certain elements of the program of studies, and it measures student ability to write under pressure. Brian comments,

> *I think that generally [the spread between diploma exam and school awarded marks] should be close, I mean some kids [marks] won't be because they can't perform under that pressure and some kids can soar under that pressure. . . . I think my class average should be higher than the exam because the exam is only testing a couple of things. You know when a student hands in a poetry portfolio and does the creative part—really works with the poem and does something really great—that is not*

> *being tested on the exam. . . . There are some very creative kids who can't write their way out of a wet paper bag and they'll score very high on the creative aspect but they'll tend to score lower on the written. And so...if that person is five to six percent higher on the school mark then they are on an exam mark then that's fine. . . .*

Brian's comments imply an important criticism of the exam's validity. Ability to write under pressure may in some contexts be important; however, there are many instances where one writes without the intense time-constraints and accompanying pressures of the diploma exam context. As well, Brian suggests that the exam only measures certain types of writing, more creative forms of text composition required within the *Senior High ELA Program of Studies* are not included in the exam requirements. On the basis of these two criticisms Brian expects that the diploma exam scores will not be a true representation of student writing ability, or student ability in relation to the program of studies. Brian elaborates:

> *I like to see [the spread between school and exam scores] close because in a way it validates my teaching, but on the other hand the exam is testing kids one day—actually over a couple of hours—in a high pressure situation in which some kids can shine and some kids can't and I don't think it always reflects a student's ability. There are some kids who need to take the time to really work things through. So really, the exam score is a good indicator of how a student might work under the pressure of post secondary education but I don't think it necessarily indicates how good a*

> *writer every student is. Some kids can write wonderfully in any situation, but some can't. . . . So yeah, the exam tests writing in a certain situation but I don't think it tells every kid this is what kind of writing you can do.*

This tension between valuing the exam in principle and being concerned about the exam in terms of its ability to measure student writing is significant. It certainly highlights one of the reasons why Brian consciously chooses not to teach directly to the test, and why he does not worry about too rigid a correlation between his students' school scores and their exam scores. Brian works through this tension by referring to a general concept in norm-referenced assessments. He states,

> *The exam is designed to (and it does it very well) bring in a certain average and if I can teach a kid to write better and read better and understand better, whatever, if I can help a student with that the exam is not a problem because that kid is automatically going to be on the right side of that average and that is why I don't think the exam drives me, drives the way that I teach because if I do my job it will work because they'll be where they are supposed to be.*

Brian minimizes his concerns over the exam's construct flaws suggesting that what the exam is specifically measuring or not measuring is not entirely relevant. His job is to teach his students to improve ~~their~~ reading, writing, thinking skills. If he does that, he feels students' scores on the exam will be above the provincial average. He elaborates,

> *I am here not to teach so that my kids are close to the exam mark. I see my role as to try and get them to develop a love for literature and poetry*

> *and prose and whatever it is to improve their writing. And I always feel I fall short.*

Brian's sense of purpose ultimately determines the major focus of his teaching. While he does design some of his instruction and writing assignments with the diploma exam in mind, he does this out of a respect for how important exam results are for his students' futures and for his professional obligation to prepare them for this exam. One might argue that Brian's focus on preparing his students for the exam is a result of its construct problems: if the exam measured writing skill more effectively, and if it reflected the curriculum more completely, Brian would likely feel less of a need to directly and consciously prepare students for the exam.

Heather's Profile

Heather is in the fifteenth year of her teaching career. She began her career in her early thirties after having completed her BEd in English. Heather is currently the English language arts department head at a large high school which is situated in a bedroom community north of a major city. Her school offers an education to almost 700 students in grades nine to twelve. Heather's students come from a mix of professional and farming families.

Heather is thoroughly committed to her work. She exudes passion for teaching, and she supports that passion with endless hours of work. She suggests that if she were to keep track of the number of hours she put into her work each week they would total more than 75 hours per week. Most weeks she works six days. Her commitment is exceptional as evidenced in her description of her teaching practice:

> *I'm a mother hen. I'm no great scholar but I work very hard at what I do. I am very interested in the literature so I study, I study a lot, and I never come into the classroom without reading the material freshly, thinking about it, thinking about it anew every single year. I don't keep lesson plans for that very reason. That doesn't mean I don't have old materials that I call on but I don't assume that I will use them, I am always looking for fresh approaches*

In addition to being committed to her work, Heather also brings other strengths to the classroom: She is a nurturer who focuses on student context, growth, and potential; she sets high expectations for her students and for herself; and she builds her practice around the relationships she forges with her students. Heather understands her students, she pushes those who need prodding, and she embraces those who need nurturing. And while her approach to teaching is focused on students as individuals, she maintains a standard expectation for them all: you will try, and you will learn.

Heather attributes much of this approach to her previous teachers and to her previous life experiences. An importance source of this experience is her mother. Heather's mother was a teacher and a nurturer. Heather recalls an experience in her mother's classroom:

> *And I can remember being appalled—there was a fellow named Opi in our class—he was bad, he had red hair and the fury to boot, and every teacher*

> *had problems with Opi. My mom came in and he was going to make mincemeat out of her and I was going to be shamed, it was going to be terrible. And he started to misbehave and she sat down beside him, and she put her arm around him and she said, "Now Howie," cause that was his real name, and she just killed him with kindness. . . If my mother has given me anything in life it is that infinite capacity for saying whatever was, was; we still have more to come and let's enjoy that.*

This sense of compassion and patience are infused throughout Heather's teaching. She understands her students on a personal level and she uses this knowledge to shape the manner in which she works with them. Heather illustrates her nurturing approach to teaching, describing her conversations with a few students who had not handed in their work before a report card deadline:

> *[One boy], he felt desperately ashamed of himself. And I said well can you just tell me why it's happening because it's getting worse and worse. He said, "because I won't do my English in the change room in hockey practice and I won't do my English as Mom's getting supper ready around me." He said, "I do English when I'm alone." He said, "I can't do English unless my soul is with me." So he said, "I haven't had time to do English." And he was being so honest. So I've made this concession and I know that what I get will be top-drawer stuff. But that person has transgressed. He's passed a line. He should be taking a zero. . . . Another one, she came and she had her papers in her hand. She said,*

> *"Will you take it, this is everything since September." [It is the end of May and] my marks are due tomorrow. And I just looked at her and I mean I'm as soft as soap....and I said, "Yes but I need a hug for it." So she gave me my hug and away she went. The girl is literally dying. She doesn't want to tell anybody. Student services have told us. I'm not supposed to know she physically is dying. She will be dead before the year is out. How could I say no to that child? How could I possibly ... It just has to be that way because if there's no relationship between them and me, then I don't have anything. I have nothing to offer but who I am.*

Heather is an exceptional teacher. The cornerstone of her pedagogy is her capacity to build effective and meaningful relationships with her students. She gives of herself to them in the expectation that they in turn will give back to her. In this spirit of giving and sharing Heather's teaching is built around discussion and dialogue. Her classroom, for example, is designed so that students sit in a horseshoe pattern. The center of the room forms a small stage area in which students can perform their work. The horseshoe pattern ensures that students are able to speak directly to and see one another. Heather chooses this arrangement because it facilitates discussion. She comments:

> *My favorite way of being is when we push the desks away, chairs in a circle, and we sit there and have discussion, sometimes a question can serve as a whole block, . . . it is more the group dynamic as learning rather than me teaching them, I am not an information disseminator.*

Heather's classroom is filled with ever-changing examples of student work. During the month I was observing her classroom, students had just completed a unit on Elie Wiesel's novel *Night* (1960); On the back wall of the classroom was a stunning mural juxtaposing key images, quotes and accompanying news stories depicting the holocaust; in the front corner of the room was a three foot by four foot model of a German concentration camp, and across the front of the room were posted a number of collages. The extensive display of student work throughout the classroom extends the dialogue by providing students an opportunity to represent their perspectives to their classmates and to others who enter the room. Heather further facilitates dialogue through the assignments she asks her students to complete and through her responses in the margins of student work.

> *[I do] a lot of freewrites, life quotes, soliloquy responses, any marginalia, responses to poems. Whenever they get an essay assignment my preamble is usually in an essay type format. They read that before they even write their own. I certainly have done multiple assignments. I'm not really big on giving them in class essays. But if I do, I write the same essay that they do. Letters, I write them letters. . . . The most important piece of writing I do is often a response to their writing which often is as long as the writing they give me. I respond philosophically to their philosophy. My marginalia round their work, that is potentially the writing that they see the most of from me to them. . . . One of the girls in her Life Quote*[1] *had*

---

[1] Heather's *Life Quotes* assignment requires students to select a number of quotes from the text that they are studying, comment both on the context surrounding each quote and on the importance of the quote for the text as a whole. They are then asked to write a piece of reflective text which explores what implications for their own lives they might draw from these quotes.

> *written about her connection to her great grandmother which of course was a profound story for me, such a beautiful connection to my grandmother, and we shared and swapped stories for a couple of papers, and it so moved me that I ended up writing easily as much as she did back to her and we became almost correspondents through this thing for two or three papers.*

Heather's writing both provides opportunity to extend her dialogue with students into their own writing, and it also provides her with an opportunity to model effective authentic writing to her students. The process certainly helps to strengthen and build relationships with her students.

Watching Heather teach, I am constantly struck by her energy. Her classroom often feels like a carnival. At times she breaks out in song, at others she encourages her students to sing, at others still she breaks into impromptu performances depicting characters in the play or novel being studied. It is difficult to predict what might happen next in Heather's room. Heather explains her unpredictability:

> *I'd say the influence that most affects me is the dynamics of the given class. At this moment, there is this class in front of me and this creative student, and this very academic student, and this very dramatic student, and this very sick child, and this one who isn't eating. Now what's going to work for them? So I try to build the program around that so I'm constantly reinventing and shifting the ground. It's very experimental I'd*

> *say. And because it's that way it's extremely labor intensive, very unpredictable and it's a lot of fun. The kids are my biggest influence*

In spite of this carnival like atmosphere, it is clear who is in control, who is managing the learning environment. Heather responds to her students' moods, needs, and interests and they respond to her with a quiet respectfulness, and a genuine engagement. Her classroom control suggests that while students are to enjoy themselves, they are also expected to perform, and to meet high standards.

Heather's teacher training program, specifically her English Methods courses further solidified her approach to teaching. While her methods courses provided Heather with a strong foundation upon which to build her practice, Heather commented on one weakness in the program: it provided no direct instruction on methods for teaching writing. Heather has made up for this in her own commitment to professional growth. In her conversations with me she discussed how writers such as Nancie Atwell and James Britton have influenced her teaching. Heather's pedagogical stance, when it comes to teaching writing, is in fact quite strong. She is teaches in the spirit of Donald Murray, as a writer who teaches writing. She claims,

> *When they write, I write. It is as simple as that. When they ask me to share from my writing I share authentically from my writing, I don't collect the best, and very often if they are writing papers I will write papers and if they are doing grading, my paper goes in the mix, and I don't always get a hundred.*

Heather uses her own writing to teach her students how to write. She demonstrates in her writing the qualities that she hopes they will develop in their own work. Heather models the features that she expects to see in her students' writing, the commitment required to write effectively, and she highlights her own commitment to writing for her students when she asks them to write for her.

Heather's writing program is constructed around a range of principles. One principle is practice. Over the course of a single school year, Heather marks up to sixty writing assignments for each student. A second principle is skill development. She accomplishes this in part through the use of a series of extensive handouts on various aspects of writing—from genres, to introductions, to vocabulary, to transitions, to process tips, to essay structure: Heather provides students with a comprehensive set of notes on writing. Heather's handouts are not the standard replicated handouts that one pulls from a teacher resource manual, or that one down loads from the internet. Each handout carries Heather's voice. For example, her handout on improving style begins with the following:

> Over the next couple of terms we will devote some of our attention to the general improvement of writing skills. Knowing that all of this material has been taught to you before, I feel no need to laboriously and/or meticulously "flog a dead horse". However, it is easy to forget the fundamentals of quality writing and focus solely on the information. Quite rightly though, an English 30 student is expected to have a sophisticated style as well as a strong appreciation of literature and literary

> conventions. That is the rigor of this scholastic venture. . . do not be faint of heart.

The strength of this approach is that it enables Heather to use these handouts not merely to convey information, but to also establish expectations for performance and to provide students with a sense of reason or purpose for developing the skills she is asking them to work on.

She also focuses on skill development through her assessment practices. Heather does not conference extensively with her students, and she does not collect multiple drafts of student work, asking students to revise and polish them. Rather, she uses her marking as a means of conferencing and as a means of encouraging students to develop their writing skills. As suggested by her earlier comments, Heather views marking as an exercise in teaching, not merely as a means to placing a number on an assignment. Heather treats each writing assignment in some sense as a draft in and of itself. Each draft builds toward developing the skills needed to write more effectively, each draft building toward preparing to write the ultimate of first drafts, the diploma exam.

Heather handles writing process in her teaching in two ways. First, she provides rich opportunities for students to engage in prewriting, pre-thinking activities. She uses freewriting as a means of generating ideas for both discussion and writing. During the time that I observed her teaching, Heather was working with students on an autobiographical poem assignment, modeled after the poem "Autobiography" by Lawrence Ferlinghetti. She began the assignment by working through the poem with the students identifying key bits of information and important elements of the poem's structure. She noted that Ferlinghetti made reference to not merely to his own experiences but he placed those experiences in the historical context surrounding his life, she pointed out allusions to popular culture of the day, and she engaged students in questions exploring the poem's purpose. After having completed this analysis she asked students to consider their own biographies, reflecting on both their personal experiences and the cultural/historical experiences of their times. In groups she had them reflect on the popular culture of their day, recalling from memory television shows, commercials, and jingles, she had them research major historical events that occurred during their life-times, and she had them focus on their immediate experiences as students and children coming of age in a small town in north-central Alberta. As they pulled their memories together she held full class discussions in which students fed off and expanded upon each other's memories. Once this was completed she had students write as many references as possible onto separate sticky notes. She then gave them two large sheets of paper, one to keep the unused ideas, and one in which to park the ideas they would be using in their poems. She had them organize their ideas using these sticky notes and then she turned them loose on their writing. From that point on her students worked independently, during their own time, on writing the poem. In terms of class-time, the process ended when the drafting began. Heather, reacting to student comments on writing process, comments on her approach to teaching writing process.

> *My first comment is "What a shame that they see [each piece] as an isolated process." Who knows the way they were internalizing the question to begin with. But I would hope that, it depends on if we are*

> *talking about the Ferlinghetti or the Life Quotes, any personal piece of writing. . . . I would be devastated to think that the community wasn't part of the process, knowing that that is so much of the preamble to the proces, of writing; Maybe they saw the process of writing as not being part of where it begins with the literature, where we usually have freewrites, prethinks, readiness for literature, the reading of the literature discussion, class discussion, group discussion, writing along the way, and then you know, the culmination, it is a much bigger process, but that's my pedagogy. . . . I am going back to my full year thing, knowing that they will have multiple opportunities that this one is a benchmark of where I am right now, and I can see by the results I get back, because I can't, David, give 60 some pieces of writing where you give first draft, I give it back with comments, you give a second draft, I give it back with comments. So that process happens between copies between different assignments, they know that it is not the end of the road, for sure not, they don't have to improve everything this time, because so many opportunities are going to be coming their way and they have openly talked about in class, "when I did this in this particular paper,' because we have had multiple conversations over the years, what do you remember pulling from other papers where it has gone well?*

Heather's perspective on drafting is an interesting one. She sees growth in student writing from one assignment to the next, and she sees that growth building toward developing the skills needed to perform effectively on the diploma exam. Within this context she does not feel a need to collect multiple drafts of individual pieces of writing. She feels that she achieves the same pedagogical ends by having students reflect on previous writing when planning their next writing assignments.

Though Heather does not work through multiple drafts with her students she does provide support to assist them in developing an effective process. One handout she prepared for her staff, titled Strategies that Work for Me outlined 14 strategies she used for helping students with their writing process. These strategies included, among others, tips on journaling, freewriting, oral response, metacognitive awareness, and chunking (organizing materials and ideas). She comments on her approach to process:

> *I think it's really important to acknowledge that they will have different processes from each other. That freewriting is not necessarily outlining, it can be outlining, mind webbing, free writing, scrambled, just as you go jotting down ideas. And then some people do it on computer, that's a different process itself but as long as there is a sense that just because it falls out of your head, that doesn't mean that that's where you will settle.*

Heather recognizes that being explicit about writing process helps students to understand that there is more to writing than putting a draft on paper, planning and revising are equally important elements of the writing process.

In addition to her focus on practice and skill development Heather builds her writing program around an emphasis on developing voice. She feels that learning to develop an effective personal voice is the key to success in writing. It is through voice that we combine ourselves with our purposes and our messages, and it is through voice that we harness our minds with our passions. Heather expresses her perspective on the importance of voice:

> *I think, first of all, honoring that voice is of paramount importance because that is personal, that you are going to shape your identity through personal nuance, and that it's important to develop that, an awareness of yourself, but it is more than just identity. . . . Identity is where the forces align that create you. Whereas the integrity is how you react to those forces. So, do they have a sense of identity, but that they do have a sense of integrity in correlation with—not just saying, "This is me and this is enough, I stop here." But, "This is me, this is what I care about, and this is how I can interface that with what you are asking of me."*

Heather's focus on voice also requires that she individualize instruction. For example, Heather's unit on *Night* is structured to maximize student individual exploration of the novel and its implications. Heather structured the unit around six stations: art, debate, Weisel and his work, parallel texts, poetry, and philosophy. Students spend three classes working in groups at these stations. Each station requires students to engage in a different activity related to the novel. The debate station for example asks students to either debate whether or not the world has learned its lesson from the holocaust, or whether or not they believe Canadians are capable of perpetrating a holocaust of their own. In her preamble to this assignment Heather writes:

> I hope you will enjoy the process of peer tutoring/elbow servicing!!! and information synthesizing. The collaboration may be the most fulfilling part of the assignment. Your ability to read directions independently and discuss a piece of literature without significant teacher intervention is meaningful learning at its best.

Heather ends this unit with an assignment titled "End of Unit Chit Chat" which provides a list of topics that students are expected to be prepared to discuss in class. While the discussion topics are teacher generated they remain focused on the students, their perspectives, and their experiences. Given the student-centered nature of Heather's English Program, the teacher directed discussion topics can be used to ensure that topics Heather thinks are important will be discussed in class. While the nature of the assignments vary across units, Heather's units share a common focus on student-centered teaching. The approach to literature supported by Heather's pedagogical stance stems from Rosenblatt's transactional theories of literary studies. And Heather's writing program is intimately tied to her literature program; virtually every writing assignment she requires her students to complete is designed as either a response to literature or is prompted by an idea contained in the literature being studied.

Heather's approach to teaching is also informed by her school context. Heather teaches in a small bedroom community. She knows many of her students' parents, runs into them at the local IGA, and regularly talks with them about their children. In this context, she claims to have "parent teacher interviews" virtually every day of the year. Her personal interactions with parents' help to build a sense of trust and respect which contributes to a certain freedom Heather enjoys in selecting materials for teaching. She comments,

> *Culturally I can pick anything I want to, they don't challenge me on that. I hope I don't ever get into a situation where I have to defend every choice that I make, I am not out to cause controversy, you just chose something, you read it, you love it, you just want to go with it, so I think I have been kind of spoiled in many respects.*

Heather's interactions within the community also support her personalized approach to teaching. As she meets with parents on a regular basis she is also able to develop a more complete understanding of her students' lives and contexts.

Yet Heather's interactions with parents are not entirely positive. Heather feels that while it is her job to nurture students, she also is required to be honest with them and their parents when they are not performing well. She has an obligation to hold them to a certain standard of performance. She finds, more now than ever before, that parents are less willing to recognize weaknesses in their children. And, she suggests that parents who do recognize areas in which their children need to grow are often unrealistic in their expectations of what a teacher can achieve in a single year. She comments:

> *Parent egos. I'd say that's a huge [societal issue]. There's a lot of bullying going on there. "What do you mean my kid isn't get 80 percent? They always get 80 percent."... The rose colored glasses and the bear claws, they all come out. And it makes it rather impossible if they don't want to....if they don't love the child enough to say I love you and let's*

> *work on this 60 to make it a 70. They just want you to fix it and when you don't they're angry at you. . . . . I do everything I can but I'm no miracle worker. I don't have a wiggly nose or a magic wand, or mystery powers.*

Part of the problem, as Heather sees it, is a systemic one. She suggests that parental bullying of teachers reaches down into the elementary end of the education system. As a result children are coddled into thinking that they are more capable than they are, that they have developed further than they really have. Heather comments,

> *[Kids] don't have to do anything to get into senior high. You have to breathe. And that's honestly it. We can't point our fingers at junior high teachers, who cannot point their fingers at elementary teachers—everybody's working their guts out—who cannot point their fingers at daycare workers. . . . There's so much potential and then society says, "No, they don't really have to do that. Don't be so mean. You're not going to keep my kid after school, not happening." And all of a sudden the standard doesn't exist. . . . If we can't be honest, if we can't say, "No you're not meeting the standard." It doesn't mean that we don't like you, it doesn't mean that you're not good enough as a person, but it means that your writing doesn't meet the standard. Maybe you don't have to redo a year, maybe you have to do intervention courses".... It keeps falling back to the person in the ditch and we're going over the top and we're in no man's land. We feel like cannon fodder. . . Like we're damned from either end. We have to meet the diploma standard.*

Being an English 30-1 teacher, Heather is acutely aware of the high standards the exam places on student ability in English language arts. She recognizes that regardless of the skills students have developed from Kindergarten through Grade 11, all students (and their English 30-1 teachers) are measured by the same rigorous standard at the end of grade 12. Her job, in part, is to prepare students for that exam. Heather also recognizes that there is pressure from her school's parent community, for her students to do well on the diploma exam.

> *Sure [the exam influences my teaching]. I think it would be negligent to say it doesn't matter. The parents and the students need to feel assured that if you want to give them this beautiful strategy they are still going to have the strategies they need for the exam, they are going to be well aware of what is going to be in there. How to do it. What to take into it. How to prepare. They've got to know that, otherwise you are saying I don't have any responsibility toward that at all. But you do, it is a professional obligation, it is a community obligation. Those parents expect that to happen, so we do make it happen.*

Given this pressure from her school community, and her sense of professional responsibility, Heather's teaching is influenced by the diploma exam. Yet Heather's views regarding the exam are, as she calls it, "multilayered."

On the one hand Heather sees value in the diploma exam program. She admires the exam manager and the personal integrity she brings to the exam program. She believes the exam operates with students' best interests in mind. She values the exam's rigor, the fact the exam provides teachers with a standard of performance to aim their students toward. She speaks of writing the exam as participating in a celebration of learning, as an event for which students have been preparing during their final year of high school. The writing produced on the diploma exam she sees as being the final draft of the writing her students have been working on all year. As they complete Heather's assignments, her students are expected to be building up perspectives, knowledge and skills that they will showcase when writing the final exam.

Yet in spite of Heather's praise for the exam she is troubled by certain aspects of it. Her most significant concern relates to the political nature of the exam program. She sees a number of facets to the exam's political nature. She is troubled by the high-stakes of the exam, by the fact that in spite of a whole year of student schooling, one exam, written in one day, has a disproportionate influence on students' futures. She is also troubled by the way in which exam results shape teachers' perceptions of themselves and of their colleagues. Heather recalls one powerful, perspective altering experience. She used to savor the first day back at school when in the first staff meeting of the year diploma exam results would be discussed.

> *It came to an end for me the year that one of my colleagues, who I think is a brilliant teacher, and who teaches science, and not only was it less than provincial average, it was significantly less than provincial average, and he was so humbled, so humiliated, that the following year on the very first day he stood up and said that he was sorry and said that is why he would no longer be teaching grade 12 biology, chemistry. . . . And it was genuine*

> *on his part, and to me I was looking at a broken man, and I was so ashamed of myself, so ashamed for feeling proud of something I didn't do, of just something that happened to be, and so ashamed to be sitting in that room, where that man who in my estimation is a quality human being with so much to offer had to be dismissed. . . And I thought, "you fool, there is pride, what difference does it make, you got your results because those kids walked in that day did that job on that day, and the rest is irrelevant, and you could be anybody, it could be you anytime, there is no guarantee of that."*

Heather is concerned that the exam is used to make judgments about the quality of teachers when the teachers have no control over how the students perform on the exam on the day that it is administered.

Heather is also troubled by the Fraser Institute's use of exam scores to rank schools. She is troubled by a ranking system that expects school and exam scores to be in close alignment with one another, and she is concerned with a ranking system that implies truths that it cannot substantiate.

> *So I think it is one of those measuring rods that really is empty, hollow, doesn't mean anything; You can make statistics say anything, which is why. . . the most important way we react is not with numbers, but it is with words, it is with real life stories, we are as a culture, we are built on story. All cultures are, so that is what we value most, we presume that the numbers insinuate this particular story and they don't necessarily.*

Additionally, Heather is also concerned about the variability within groups of students: in smaller schools especially, a handful of students can change a school's results dramatically. From year to year, without any changes in teaching practice a school's results can climb or fall simply in the basis of variance in student population.

Given these concerns regarding the diploma exam, Heather attempts a balanced approach in regards to preparing students for the exam. With her own students she attempts to de-emphasize the exam. Rather than have them focus on the exam during the course, she has them focus on learning, on developing voice. She explains:

> *I tell them, "You're not here for the diploma. I promise you, you will pass that diploma. I've never had anybody fail it. You will pass it, I promise you that but you have to pass my course. So let's talk English, let's just build a community of learners. There's no magic in that. It's just that your language will change, the way you talk about it changes, the focus is not on writing is for diploma exams."*

Heather does not want the exam to become the focus of her course. She wants her students to focus on skill development and enjoyment of the material. With this in mind, Heather structures her course to minimize its focus on the exam. Rather than spend class time on exam preparations her students are given multiple choice reading tests to take home, and in the spring when students are preparing for the exam she schedules four exam preparation seminars before school begins, from 7:15 to 8:15.

Heather sees the exam as compatible with her approach to teaching. Teach the literature, know it intimately, have something to say about it and when it comes time to write the exam you will be prepared for it. Heather's approach to writing process, too is quite similar to that of the exam. She spends much time on developing ideas and voice and prewriting, but does not focus within individual assignments on revision with her students. In many ways her pedagogy, though perhaps not directly influenced by the exam, aligns with the construct the exam is measuring. Heather explains the alignment as follows:

> *So of course they are diploma preparing all the way along, but it isn't labeled as that, and it isn't directed at that test until that last go round. "Now bring your literature back and decide what in the mix was most passionate for you to write about, what do you really feel that you have something important to write about, now choose some." "What is some?" "I don't know. If you are a huge* Hamlet *fan and you are passionate about it, you still have to be prepared that if the question doesn't match nicely, dovetail nicely, or if you get in there and you go blank, you need to have something else to access." So they need to do that. But it is not an information driven course, it's for skills, so that is what we do; We learn about our skills, we hone our skills, it is an art course.*

Over the years Heather has developed her sense of pedagogy in relation to English language arts. She understands what the curriculum is calling for—skill developments rather than explicit information regarding a text—and she understands what the exam is measuring. Additionally she has a firm sense of her identity as a teacher and what that implies for how she will teach English language arts. As this sense of pedagogy developed, Heather's focus on the exam became less pronounced. But this shift away from focusing on the exam was many years in the making. Heather comments:

> *The first year I taught English 30, I had no idea what was even in the exam, and because I came from Nova Scotia, I had never written a diploma; I didn't know what the word insinuated, so I thought "Oh, okay, I'd better find out about that." So I did. And I am sure that I dealt with it in some mediocre manner, and they wrote and they did what they did and the results from that year are probably going to be pretty much replicated by the results of this year. It just won't change. . . . I think I have built up in the middle of my thus far career, a stronger emphasis on diploma testing . . . Once a term, or maybe even once a unit, I can't recall, I'd bring out the dog and pony show—let's do a little dance around this and write one, and "Oh you're getting better, you are getting better." They still do the same at the end of the year that they did at the beginning;, really it is personal progress that it is measuring. And so over the last five years, [my focus on the exam] has diminished, diminished, diminished. . . . Even as late as last year though I would have done probably one a term. . . . This year for the first time ever I didn't have a mid-term or a final, and I thought when I made the decision that come the end of the year I would*

> *feel really nervous about that. I don't. It wouldn't have made any difference anyway, probably would have just added to the stress.*

Ultimately, however, in spite of Heather's shift away from the exam in terms of her own pedagogical focus, the exam's influence can be seen in her writing program. Many of her writing assignments reflect the types of writing required on the diploma exam while a number of writing assignments administered throughout the year are modeled on previous writing exams. Heather's exam preparation work stems from a sense of professional responsibility toward her students rather than from a philosophical affinity with the purposes and goals of the diploma exam program. She comments:

> *For some folks that idea of teaching to the test gives them a method and a reason. And that is a good thing for them. For me it is not. Philosophically I want to discredit that I am a political vehicle. That is probably just a personal angst that I don't want to be pigeon holed like that. I'd like to think that there is something more noble in the whole thing than just preparing for a test in case you need the marks to go to university. It seems so narrow, it seems so small thinking*

Heather's focus is on the larger picture on preparing students for life rather than merely preparing them for an exam, yet she also recognizes that at times preparing for an exam is part of preparing for life.

Analysis Across Cases

Stake (2005) argues that in multiple case study research, it is important for the researcher to first focus on each case independently from the other cases under investigation, to develop a thick description of each case, to look at what information each case has to offer in relation to the issue under investigation. Once this has been completed a discussion across cases can take place. The following section contains an analysis across cases, focusing on both the similarities and differences between cases. Collectively this analysis is designed to explore common experiences in teaching English 30-1 while also describing how differences in perspective and context shape teachers' approaches to pedagogy. This analysis focuses on these issues with a primary interest in each teacher's transaction with the English 30-1 diploma exam and its impact on teaching and learning of writing.

Sources of Pedagogy

The teachers involved in this study approach their classrooms with a concern for their students as individuals, with a well developed sense of professional responsibility and with a desire to serve their students well. Each teacher emphasizes the value of developing and maintaining open and respectful relationships with their students, and they use their relationship building skills to enhance their students' classroom experiences. The pedagogy displayed by each of the teachers participating in this study stems from a range of sources which include the program of studies, the diploma exam, previous educational experiences, and teacher training programs.

Senior High ELA Program of Studies (Program of Studies)

The Program of Studies is the official government-mandated curriculum for senior high English language arts. Legally, teachers are required to plan their courses and their instruction to ensure that students are able meet the outcomes required by this program. It is understandable, then, that this Program of Studies exerts significant influence on the pedagogy of the teachers participating in this study. Key features of pedagogy stemming from the Program of Studies include the development of units which involve the integration of the six language arts strands—speaking, listening, reading, writing, representing and viewing—into a common set of assignments; the focus on writing as a response to literature; the emphasis on metacognitive thinking skills; and the use of representation as a means to extend and develop one's thinking.

Previous Educational Experiences

Madeline Grumet (1992) argues that our previous experiences as students influence our pedagogy in the present. Her argument holds true for all three participants in this study. Brian's high school experiences cause him to value the diploma exam as a means for holding teachers accountable to the curriculum. His experience with a junior high teacher prompted him to use humor as a means of relating to students. Heather's elementary and undergraduate experience reinforce for her the importance of being a teacher who nurtures, it suggests to her the value of high standards, and it points her to dialogue and discussion as an important means of learning. Anne's previous educational experience teaches her the value of both believing in students and in pushing them to achieve.

Teacher Education Program

Teacher training programs also have an important influence on each teacher's pedagogy. Heather describes her experiences in a rich teacher education program, one in which she was challenged to develop her personal philosophy of education and come to understand how that philosophy shapes one's approach to teaching. Her methods course, while demanding, focused on reading and reader response while neglecting to teach writing pedagogy. She comments:

> *I never had five minutes worth of instruction [in how to teach writing]. Zippo. So I mean I'd be interested because I've never heard it taught to me what I should do to teach kids writing. Isn't that funny? What a strange thing. Funny. Peculiar.*

Heather attempts to address this deficiency by reading from some of the leading figures in the field; however, it is understandable that, given this history, her pedagogy related to the study of literature and reading—she is student-centered, discussion-oriented, and transactionally-based—is stronger than her pedagogy related to writing. While Heather is process oriented in her teaching of writing, she focuses primarily on the front-end of process rather than the back-end. Similarly, while the marginalia she composes focuses on her personal responses to student text rather than merely the correction of error, literature regarding teacher comments to student writing suggests that frequently students either do not read comments in the margins or that they frequently misunderstand such comments (Zellermayer, 1989). Zellermayer demonstrates in his review of literature on feedback to student writing, that written comments on student writing is the least efficient and least effective method of responding to student writing. As well, Heather's contention that students grow between papers as they apply knowledge gained from one paper to the next, while understandable, is also problematic. Growth in writing stems from applying lessons learned as they are being learned. While students certainly can transfer learning from one paper to the next, research also suggest that if learning is transmitted through written comments on previous papers, students rarely apply these comments to subsequent writing.

Brian's teacher education program, too, left Brian unprepared for teaching writing. Brian was an English minor and a Math major. As an English minor, Brian would have taken a small number of English courses:

> *Math is my major and I graduated with an English minor. I didn't even have an English practicum. I took English courses dealing with literature. I took one that dealt with linguistics. Things like that. I never took a writing course in my degree.*

Brian's English methods course ran for sixteen two and a half hour classes. While Brian certainly would have been exposed to some instruction on how to teach writing, he suggests that this was quite limited. Generally speaking, at the institution in which Brian completed his BEd, each student completes a practicum in his or her minor. For one reason or another, Brian did not receive a practicum in an English classroom. Prior to teaching English in his own classroom, Brian had little formal experience in teaching English. As a consequence of the limitations in his teacher education program, Brian has felt insecure about his ability to teach English language arts throughout his career. Brian has dealt with this insecurity by being active in his professional development. Through out much of his career, Brian has attended most of the annual ELAC conferences. His practice certainly reflects this. A number of his key assignments have been adapted from presentations he has attended: his found poetry assignment, poetry portfolio, numerous representing and metacognition assignments all stem from ELAC sessions he has attended. In terms of his work with literature, Brian is on the cutting edge of reader-response approaches to teaching. Brian's teaching of writing, however, has not achieved the same growth. In part, this is due to the fact that Brian has attended few sessions on writing at ELAC conferences. Brian, has attempted to attend such sessions but has found them to be few and far between and frequently focused on grade levels below the ones he is teaching. Issues related to Brian's pedagogy are related to those of Heather's. While he is process oriented, he is focused on the front-end of process, working with students on prewriting, planning and drafting, while providing limited time or focus on revision. Brain also values the formulaic structure of the five paragraph essay because it enables students to hang their ideas on a structure that they are familiar with. This focus on attending to a formula, however, frequently detracts from a focus on the development and exploration of ideas. Brian has also made an interesting decision regarding timelines for student projects. He has been frustrated that students do not use their time appropriately when completing assignments; too often they leave their writing until the last minute. Rather than develop strategies to encourage students to engage in a more robust writing process, Brian has opted to shortening timelines, allowing students less time to engage in process than they might otherwise have had.

Anne's teacher education program included a significant focus on methods for teaching writing. Anne recalls studying writing process "almost to death" and working in a junior high school assisting students in their writing. Anne feels that she was well prepared to teach writing.

> *I do remember though, that we did a lot of studying of how to teach writing, we did a lot of looking at different ways of teaching writing, we actually worked really closely with a junior high school, so that part I felt pretty confident about. . . . A huge component of what we did was writing process and teaching this and going into junior high and elementary schools and watching teachers teaching writing and evaluating . . . . [we] sort of beat the writing process to death—you know at a certain point you just say, "let it go," because the process is different for everybody, some of us plan more than others, and I think students as a whole don't plan at all. . . . I sort of figured that I knew how to teach writing.*

Her approach to pedagogy at the beginning of her career reflects her preparation. She engaged students in interesting student-centered transactional responses to literature and she developed a number of representing assignments for students to complete, her writing assignments flowed from these representing tasks. However, her experience with diploma exam results during her first year caused her to seriously question her approach to teaching. She decided that in order to ensure alignment of school and exam scores, she needed to teach explicitly to the diploma exam. Therefore she decided to abandon much of her progressive approach to teaching in favor of a transmission approach to teaching literature and a formulaic approach to teaching writing. Within this shift, however, she maintained a focus on peer and teacher conferencing.

Diploma Exam Experiences

Each of the teachers acknowledges the importance of basing their instruction around the requirements of the English 30-1 Program of Studies while simultaneously ensuring that students are prepared to perform well on the English 30-1 diploma exam. They recognize, therefore that their teaching practice is influenced to varying degrees by the diploma exam. While differences exist in terms of the exam's influence on each teacher's practice, several key similarities across the three teachers' practices are evident. These similarities include the following:

- An overwhelming preponderance of each teacher's writing assignments are essay writing assignments that focus on students' responses to literature.
- Each teacher is committed to process-oriented approaches to writing instruction, though in class, actual focus and class time are limited to prewriting, idea development, organization and drafting. In each classroom time for revision, polishing, or publishing elements of process are significantly limited.
- With varying degrees of intensity, each teacher focuses some instruction on helping students develop the skills needed to deconstruct diploma exam questions.
- Each teacher relies on the diploma exam rubric when marking classroom writing assignments.

Narrowing writing focus. In James Britton's review of teaching writing in England in the 1960s Britton observed that writing assignments in courses where students were preparing for external examinations were more narrowly focused in terms of the types of writing assignments they required of students and in terms of the audience they were required to write for. He summarizes his observations as follows:

> From years one to five the chief movements seem to be (a) from report to analogic within the transactional, and (b) the concentration of most transactional writing into the 'pupil to examiner' audience. Then in the seventh year, while there is some extension into the higher transactional categories, the landslide is into the analogic, and principally at the expense of writing in the poetic. This pattern, it seems to us, is most consistent with the interpretation that sees the mounting effect of examination demands as dominant (Britton et al, 1975, p 195).

Essentially, Britton argues that external exams shape the writing experiences of students.

The evidence provided in the previous case study descriptions suggest that a similar narrowing of focus and audience occurs within the three English 30-1 programs described in this study. Virtually every assignment Heather, Anne and Brian provided to me for analysis contextualized the writer-reader relationship as being between student and teacher, a structure which replicates the diploma exam's structure. Often students were given freedom in terms of choosing the topic they wished to write on or the perspective they wished to write through. Yet in each case it was the teacher who received the paper for marking. Anne did create some assignments where peers evaluated each other's writing but even in these cases the final marker, the one that truly mattered, was the teacher. Other exceptions include Heather's Ferilenghetti poem which was intended for a broader audience of parents, grand-parents, siblings, and classmates, though the primary audience for this piece was the student writer him/her self. Heather hoped that students would gain an appreciation of their past experience as they complete the assignment. So, while the poem was graded by Heather, every student who completed the assignment received a perfect grade (though students did not know this prior to handing the assignments in).

The vast majority of writing assignments students received from the three teachers involved in this study were focused on essay writing, and a majority of these focused on personal and critical responses to literature assignments. These trends certainly reflect the findings of Britton's study and they suggest that Alberta's English 30-1 diploma exam focuses (in the three cases observed) teachers and students on the writing context—essay focus and implied audience—defined by the exam. Additionally, all three teachers designed their instruction to focus students on developing skills needed to deconstruct and respond to diploma exam questions. Their approaches to this varied somewhat. All three teachers provided examples of previous diploma exam questions to their students. Brian and Anne ask students to write formal essays in response to these questions as they relate to specific pieces of literature being studied while Heather provides sample questions to her students during her early morning exam preparation tutorials. Brian explains his rationale for designing assignments with an explicit exam preparation focus:

> *I don't do a lot of real exam prep and I probably should do more just to get them comfortable. I do take exam questions from the past and I'll get them to write an essay. . . . So that way they are at least comfortable with those types of questions. I could probably do more, taking a few questions from over the years and having them. . . at least writing introductions that address those questions. So it is something that I think I have to put a little more emphasis on. . . . It is good to get them comfortable with reading a question, deciphering it and saying ok I can use this or I can use that.*

Brian's primary goal for using previous diploma exam questions in his teaching is to help students become comfortable with trying to understand both the structure and the diction contained in the exam questions, a skill that will help them more efficiently determine how to respond to the question. Anne too suggests that decoding skills are important. She suspects that her students' struggles with the diploma exam stem from their initial inability to deconstruct the exam question. The frustration generated in the process of attempting to understand the question elevates their stress levels, she argues, causing them to perform less well than they should have.

The practice of actively teaching students to deconstruct previous exam questions is often criticized in the assessment literature as being an unethical teaching practice, one that pollutes test scores (Volante, 2006). However, it is clear in the case of the three teachers participating in this study, that the motivation behind this practice is not driven by a desire to artificially inflate their students' test scores, but rather by a desire to ensure their students are able to understand a set of questions that are often written using a syntax and a diction that is unfamiliar to many students. Teaching students how to decode previous diploma exam questions is consistent with their professional commitment to ensure that their students are well prepared to write this high-stakes exam.

Narrowing the Marking Scheme. All three teachers involved in this study derive their marking schemes from the diploma exam rubric. In many cases they use the rubric exactly as it has been developed by Alberta Education. The rationale for this choice seems to be two fold. On the one, hand using the diploma exam rubric enables students to get a sense of the scoring criteria developed for the diploma exam and other hand it enables the teacher the opportunity to ensure to that exam scores and school scores reach some degree of alignment. Linda Mabry (1999) argues that in high-stakes writing assessment contexts it is understandable that teachers use provincial exam scoring guides to inform instruction and to explicitly prepare students for the test. She cautions however that standardized rubrics in high-stakes testing contexts are "overwhelming the writing curriculum" (p 676). She argues that the rubrics receive unprecedented priority in classrooms so that they become the focus of instruction and of student writing. Mabry further cautions that a consistent application of standardized exam rubrics on student classroom writing may disadvantage students who are creative in ways the rubric was not designed to anticipate. Brian recognized this issue in his own students. During our first interview Brian expressed concern regarding one of his students who was very creative. He was concerned that the kind of writing he was requiring in English 30-1 would not adequately enable this student to demonstrate the full range of her writing skills, and he was worried that she would perform poorly on the diploma exam because the rubric did not seem to value the writing genres in which she excelled.

Limiting process. The more striking issue related to the exam's influence on teaching writing emerges in the context of writing process. While Brian, Heather and Anne described their understanding of writing process to include planning, drafting, revising, and polishing, it is clear, through interviews, observations, and writing assignment designs, that they focus mostly in their instruction and in the class time they make available for completing assignments, on students' ability to engage in the beginning stages of writing process rather than the final stages of revision and polish. Significantly, the elements of process the teachers focus on are the elements emphasized by the diploma exam, while the elements neglected by the three teachers are elements that are also neglected by the exam. So the process advocated in the classroom is similar to the process required by the exam. Conforming to the exam process in the classroom assignment does not appear to stem from a direct attempt to teach to the test; rather it seems to be a consequence of the time constraints teachers feel, either in regards to their time to mark multiple drafts and revisions, or their providing class time to get students started on an assignment while not having time set aside toward the end for final stages of process to be worked on. It may also stem from students' tendency to begin assignments with too little time to engage in process in a significant way. However, it is clear that this approach to process is certainly supported by, even enabled by, an exam that does not measure or allow for writing process to occur in a fulsome manner. Given Brian, Anne and Heather's commitment to preparing students for the diploma exam it would be likely, that had the exam required students to engage in a more fulsome process, these teachers would have ensured that their students were required to complete classroom assignments utilizing a more complete process.

Context and Experience

While similarities exist across the three teachers in terms of how they prepare students for the diploma exam, this practice also exists along a continuum of intensity and directness. Anne was most direct, explicit, and focused on preparing students for the diploma exam, Brian was much less focused on the exam, and far less explicit in his preparation of students for the exam; however, a significant portion of his writing assignments was designed based on diploma exam questions; Heather was least focused on the exam in terms of her teaching, in fact she actively encouraged her students to not worry about exam preparation activities. Heather's assignments were not directly modeled after the diploma exam assignments though they did certainly reflect the type of thinking and writing being measured by the exam.

Further analysis of data suggests several possible explanations to explain why teachers might fall across various positions on this continuum.

Attitudes toward the exam. One explanation has to do with teachers' attitudes toward the diploma exam itself. The three teachers value the exam for similar reasons: they appreciated its ability to create a level playing field for students applying to universities from schools across the province. Because all students write the same exam under similar conditions, the three teachers feel that they allow for a more fair comparison of students for purposes of university admissions. The teachers were concerned that in the absence of diploma exam programs grade inflation in some schools would make it difficult for students in schools with more rigorous standards to gain acceptance to universities. The teachers also appreciated the exam's ability to establish standards, standards to aim for, standards against which teachers and students are able to compare classroom scores. Yet, in spite of these positive attitudes toward the diploma exam the teachers also express a range of concerns regarding the exam. While all three value the exam for its ability to provide information against which they could compare their school awarded marks, the teachers express concerns that these exam results are being used to make judgments about schools and teachers. Primarily this concern revolves around construct issues. The exam's construct focuses on student writing ability, not teacher competence; therefore they believed it would be improper to use test scores to make judgments about teacher quality. Anne and Heather both argue that average student performance on the exam is more a function of population than anything else. Anne argues that it is not fair to judge a teacher based on the population that he or she is working with.

Questioning the exam's construct. Further concerns revolve around construct issues. Anne is most articulate and forceful in her criticism of the exam's construct. She identifies several elements of construct irrelevant-variance and construct underrepresentation in the exam. She argues that the exam measures student writing ability in relation to a narrow set of genres: student ability to generate personal and critical responses to text, most often essay response. She points out that the Program of Studies calls for a much broader range of writing to be performed in the course including poetic, narrative, and functional writing. She also observes that given the exam design, student ability to write under pressure and within time constraints is being measured. She notes that this is an artificial and unrealistic context for writing, one that does not reflect what happens in authentic writing contexts. Brian also voices these criticisms though he is more moderate in his critique. Heather on the other hand is not too troubled by the exam's construct. In discussion with her students she makes it clear that if they follow the course in the manner she has designed it, they will be well prepared for the exam. Her comments suggest that her pedagogical approach naturally fits with the exam's construct.

Interestingly then, Heather, who is least critical of the exam's construct is also least focused on direct or explicit test preparation work in her English 30-1 class, while Anne who is most critical of the exam's construct is also the most focused of the three teachers on directly and explicitly teaching to the exam. Brian sits in the middle in relation to both criticism and practice. Based on this dual distribution on both the teaching to the test and the exam critique continua I would hypothesize that the more aware a teacher is of construct flaws in an exam, the more likely the teacher will focus on explicitly preparing students for that exam. Given the high stakes nature of the exam, each of the teachers participating in this study clearly articulates an understanding that they had a professional obligation to prepare students for the exam.

External pressures. In relation to this professional obligation, these three teachers also commented on the expectations of students, parents and administrators that students were both well prepared for the exam and that their exam scores would correlate with their school scores. Again, the pressure Anne felt in this regard was significantly greater than that of Brian and Heather. She described meetings with her administrative team and comments from parents which made these expectations abundantly clear. The pressure Heather felt in relation to marks stemmed more from some parents resenting her high standards and expectations rather than direct pressure related to exam performance. Her school administrators pressured teachers to ensure strong exam performance and alignment of school and exam scores more through indirect than direct means: department heads reported during the September staff meeting on previous year exam scores, and cakes were brought in to the staff meeting to celebrate success on diploma exams. Brian described feeling little to limited pressure regarding student performance on the exam.

The pressure to align exam scores and school scores seems to have further served to intensify Anne's focus on directly preparing students for the exam. Anne comments on this, claiming that given the construct problems with the exam, the only way to ensure alignment of test scores and school scores is to narrow one's teaching so that it focuses on the exam. Ironically, this is what Heather has done in her own practice, though not explicitly, her confidence in her students' ability to achieve school scores close to exam scores is based on the fact that the constructs upon which the school scores are based is (given Heather's approach to teaching) more naturally aligned with the diploma construct. Anne on the other hand describes how in her first year of teaching English 30-1 she engaged her students in a wide range of activities that, while valuable educationally, were difficult to assess and were not necessarily consistent with what the exam was measuring. After submitting term marks that were on average 20% higher than the diploma exam marks she decided that she needed to dramatically alter her pedagogical stance in order to ensure a greater alignment of school and exam scores. This experience caused Anne to question her professional competence. After aligning her teaching with the diploma exam she had the opportunity to mark diploma exams for Alberta Education. This experience helped to regain her confidence; she walked away from that experience believing that her new approach was enabling her to align her standards with the diploma exam's standards.

Teaching experience. A further factor which seems to impact the degree to which the exam is influencing pedagogy seems to fall along a continuum of teaching experience. Heather is least focused on teaching to the exam, yet she describes how during her career this degree of emphasis has changed. In the beginning she was interested in finding out what the exam was all about and so she studied the exam and had students do exam preparation activities: around years seven and eight of her career she claims to have focused most explicitly and concretely on preparing students for the exam, and then from that time to the present (year fifteen) her explicit focus on the exam has dissipated to the point where she now discourages students from being too focused on the exam during the year. Anne and Brian are both in year eight of their teaching careers, though Brian has been teaching English 30-1 for about five years longer than Anne has. Anne is most explicitly focused on the exam, Brian far less so. This seems to be consistent with Heather's experiences

Caught Between Paradigms

The case study data certainly suggests that the English 30-1 diploma exam is impacting the teaching of writing. Teachers narrowed their teaching of writing to focus on the genres, the audience, the marking scheme, and the writing process defined by the diploma exam context. While the exam's influence seems to be consistent across the three teachers, the degree to which teachers directly teach to the test is more varied. The more aware a teacher is of the exam's construct problems, the more directly that teacher seems to focus on teaching to the test, while the less concerned the teacher seems to be about the exam's construct the less explicitly focused on teaching to the test the teacher seems to be. Pressure from school community (parents, students, and administrators) to ensure that school and exam marks are close to one another also seemed to strengthen the impact of the exam on teaching, whereas teachers who do not feel explicit pressure to align scores do not seem to focus as much on explicitly teaching to the exam.

Assessment literature (Crundwell, 2005) suggests that explicit teaching to the test is a significant issue in terms of a test's impact on pedagogy. From a testing perspective, this may be true. From a curriculum perspective, however, it seems that the more subtle influences are an equally significant problem. In terms of writing process, the exam supports poor pedagogy by requiring students to develop skills related to a minimalist writing process. Brian, Anne and Heather conform in their teaching practice to minimalist forms of process, not out of a desire to explicitly teach to the test, but because the test does not challenge them or their students to engage in a more fulsome process. In fact, the analysis of the diploma exam in Chapter Three demonstrates that the exam conceives writing in terms of the Current Traditional Paradigm rather than the Process Oriented paradigm that emerged through the 1970's and beyond. The significance of this exam's influence on pedagogy is readily demonstrated when we consider the extent to which each of the teachers involved in this study is trapped between paradigms in terms of their writing pedagogy:

All three teachers, while they acknowledge the importance of writing process, and while they provide opportunity for students to engage in a process oriented approach to writing, rarely engage with students in their processes. When they are involved in student processes it is almost always with a focus on the prewriting, planning, and initial drafting elements of processes. Their work with writing processes does not explicitly acknowledge that writing processes are recursive. They collect final work, not work in progress and so do not often take opportunities to encourage revision as a means of reconceptualizing or expanding ideas. While Heather's assignments (e.g. Life Quotes) are designed to encourage students to use writing as a means of discovery, Brian and, especially, Anne encourage students to develop their ideas before they write and so see writing as a means to communicate developed ideas rather than as a process through which ideas are developed. In the same vein, all three teachers seem less concerned about purpose and occasion than they do about structure and audience; in their writing assignments audience is almost always defined as the teacher. Similarly, while all three teachers require their students to write in a number of modes, their overwhelming focus is on expository writing. All three teachers believe that writing is a skill that can be taught, but at times seem at a loss to explain exactly how this is done. Only Heather writes with her students. And in their assessment of student writing they focus on the product: idea development, structure, style, and usage. Each of these practices stems from the Current Traditional Paradigm of writing instruction, and each practice also reflects the construct and process defined within the English 30-1 diploma exam.

CHAPTER 6: TWELFTH GRADE STUDENT WRITERS

Nothing you write, if you hope to be any good,
will ever come out as you first hoped.
-Lillian Hellman

Not that the story need be long,
but it will take it a long while to make it short.
-Henry David Thoreau

Studying the impact of writing assessment on pedagogy is an important first step in understanding the consequences stemming from the implementation of an assessment program. A second, equally important question, however, looms behind the first, "What impact is this exam having on student writers?"

This chapter focuses on exploring answers to this question. It opens with a description of the methods used to collect student data and moves into a discussion of students' classroom and assessment experiences, and of the lessons about writing that they have learned as a result of these experiences.

Methods and Analysis

In attempting to explore the question articulated above I engaged in two methods of data collection. First, I interviewed ten students from Brian, Heather, and Anne's classrooms (three or four from each). During the case study research I asked each teacher to identify a student that he or she thought was a strong writer, an average writer, and a weak writer. These students were later interviewed regarding their attitudes toward writing, their perspectives on what was being valued in their writing (by themselves, their teachers, and the diploma exam), and the process they engaged in when writing. Student interviews were conducted using a semi-structured interview process. While the protocol questions were followed, order of questions and degree of probing varied from interview to interview. Each interview was fully transcribed. Student responses were analyzed according to two main questions:

1. What is the student's perception of what was being valued in his or her writing
    a. by him/herself?
    b. by his/her teacher?
    c. by the diploma exam?
2. What process did each student follow in his or her writing
    a. on classroom assignments?
    b. on the diploma exam?

Patterns across groups of students (by teacher affiliation or by perceived ability level) were then analyzed. The results of this analysis are reported below.

The second approach to investigating the impact of the diploma exam on student writing involved the development, administration, and analysis of a survey that was administered to English 30-1 students across Alberta. The instrument used for this survey was developed specifically for this study; it was designed to answer several questions (for complete survey, see Appendix 3):

1. What are students' attitudes and beliefs regarding writing instruction and assessment?

2. What are students' perspectives regarding their classroom experiences with writing instruction and diploma exam preparation?
3. What are students' actual practices when writing in English 30-1?
4. How do students interpret the constructs underlying writing assessment tools used in English 30-1?
   a. in classroom assessments?
   b. on diploma exam assessments?

I developed survey items and accompanying scales. I then discussed a first draft of this survey with a former colleague, an English 30-1 teacher who explained his interpretation of what each item was asking. Based on this discussion many items were revised. The survey was then piloted with a single group of fifteen students. Students' written responses were then analyzed. Unexpected variance on a number of items suggested that students were interpreting items differently from one another. Based on this analysis I decided to engage a group of students in a focus-group discussion around each item. The five students who participated in this focus-group represented a range of writing abilities as identified by their teacher. Each item was discussed using the following format: Students were asked what they interpreted each question to be asking, they responded. Students received an explanation of how each item was intended to be interpreted; we then discussed first whether or not the item was being interpreted as intended and then, second, on how the item might be modified to ensure that it was being interpreted in the manner hoped for. We workshopped each item until consensus was reached on the question was actually asking.

I then administered the survey to English 30-1 classes in seven schools (three urban and four rural) in Alberta. One hundred and thirty-two students completed the survey (a 54% completion rate).

Survey sample aligns relatively well with the population of English 30-1 students (as reported by the province) on three descriptive characteristics: gender, average school awarded mark, and average diploma exam mark (see Table 6.1).

Table 6.1. Comparison of Sample and Population data (in percentages)

	Sample	English 30-1 Population*
Male	33	44
Female	67	56
Average school awarded mark in English 30-1 **	74	72
Average English 30-1 diploma exam mark **	70	67

* as reported in Alberta Education's *English Language Arts 30-1 Diploma Examination School Report June 2006.*

** sample results relate to students' expected scores (school year had not yet been completed) while population results relate to actual scores.

Classroom Experiences of Survey Respondents

Students' responses to the survey items regarding their classroom experiences reflect the conclusions made in regards to Brian, Heather, and Anne's classes as reported in the previous section: these classrooms too, seem to reflect a situation in which composition pedagogy is caught between paradigms, though largely pressing toward the current traditional paradigm.

The students participating in this survey believed that their teachers were quite focused on the exam when teaching writing. Seventy-two percent of students either agreed or strongly agreed to the statement that their teacher focused the English 30-1 course on diploma exam preparation. In these same classrooms, fifty-five percent of students also claimed that they often or always received writing instruction through diploma exam questions. For these fifty-five percent of students, virtually all their writing assignments are either previous diploma exam questions, or they are questions that have been modeled after the exam. A similar fifty-five percent of students reported being required to either often or always structure their essays according to a five paragraph formulaic structure. These students report classroom contexts similar to what students in Anne's classes would have experienced. This finding reflects the findings of Britton's (1975) work in England, where he discovered that the nearer one got (in terms of grade level) to the actual standardized exam the more one's teachers focused on that exam in their instruction.

Table 6.2. Intensity of Student Experiences with Various Teaching Strategies in English 30-1

Variable	*Always	*Often	*Sometimes	*Seldom	*Never	Mean	Standard Deviation
Receive Written Feedback (n=132)	28.0	50.8	19.7	1.5	0	4.05	0.73
Writing Through Practice Diploma Exam Questions (n=130)	18.5	36.2	29.2	8.5	7.7	3.49	1.12
Five paragraph essay expected (n=129)	23.3	31.0	25.6	10.1	10.1	3.47	1.23
Conference with Classmates (n=132)	9.8	23.5	31.8	22.0	12.9	2.95	1.17
Conference with Teacher (n=132)	5.3	17.4	33.3	32.6	11.4	2.73	1.04

* Responses in percentages.

Feedback to student writing also appeared to be grounded in approaches used which reflect the current traditional paradigm. Seventy-nine percent of students claimed that they often or always received feedback in written form. On the other hand, only twenty-three percent of students claimed to have often (17.4%) or always (5.3%)

conferenced with their teacher about their writing. Conferencing with classmates was more common (33% often or always) than conferencing with teachers.

What Students Value in Their Own Writing

Eight of the ten students interviewed valued writing that was personally meaningful, that dealt with ideas and experiences that were important in life, and that enabled them to develop new ideas. The students who completed the survey too valued writing that was personally meaningful to them. Ninety-three percent of them believed that they performed best on personally relevant writing assignments. Eighty-five percent of them also believed that an effective writing process was important while sixty-eight percent believed that reshaping one's text was an important element of effective writing processes. Unfortunately, their understanding of effective writing processes appears to be limited: Eighty-six percent believed that clarity of ideas was an important goal of first draft writing; sixty-one percent believed that clarity of expression is an important goal of first draft writing; fifty-two percent believed that it was important for a first draft to be organized.

Table 6.3. Students' Beliefs about Good Writing

Variable	*Strongly Agree	*Agree	*Neutral	*Dis-agree	*Strongly disagree	Mean	Standard Deviation
Write Best on Personally Meaningful Assignment (n=132)	74.2	18.9	4.5	2.3	0	4.65	0.68
Effective Process Foundation of Good Writing (n=132)	25.8	59.1	12.9	2.3	0	4.08	0.688
Clarity of Ideas First Draft Goal (n=132)	25.0	60.6	7.6	5.3	1.5	4.02	0.82
Reshaping an Important Element of Process (n=132)	16.7	52.3	27.3	3.8	0	3.81	0.74
Clarity of Expression First Draft Goal (n=132)	14.4	47.0	22.0	15.2	1.5	3.57	0.97
First Draft Organized (n=132)	9.8	42.4	25.8	18.9	3.0	3.37	0.99

* Responses in percentages

The effect of the diploma exam rubric on writing instruction is revealed quite dramatically in students' responses to questions regarding their perceptions of what their teachers are valuing in their assignments and assessment of student writing. For example, of all the students interviewed, only Heather's students believed that their teacher valued writing that was personally meaningful. Brian's students believed that he valued writing in which ideas were well defended and well structured while also valuing correct sentence structures and effective use of diction. Anne's students agreed less on what she valued in writing: her weak writer believed that she valued spelling and grammar, while her average writer believed that she valued writing which is focused and on topic, while her strong writer believed that she valued ideas, support, structure, and stylistic choices.

Student Analysis of Classroom Writing Assessments

One section of the survey asked students to answer the following question: "How important are the following skills for successfully completing and receiving high marks on your classroom writing assignments? Your ability to:" followed by fifteen skills which fit into one of three general categories: explicitly stated elements of the diploma exam construct; unstated construct elements which introduce irrelevant variance into the exam; and elements of the construct which are not being measured by the exam. Skills which are directly stated in the exam rubric received the highest responses (see table 6.4).

Over eighty percent of students believed that each of these skill sets was either important or very important in achieving high scores on their in-class assignments in their

English 30-1 classes. The skill sets highlighted by students each appear explicitly on the diploma exam rubric. These skill sets also correspond to features of text—diction, syntax, grammar, mechanics, organization—and focus on the correctness of student writing. As such student responses to this section support their responses to the section in which they describe the context in which they were learning to write. As with Brian and Anne's students' interview responses, these survey responses suggested that their teachers are largely locked in the Current Traditional Paradigm in terms of their writing instruction. The emphasis on elements of correctness in the assessment of student writing assignments further reinforces this picture.

Table 6.4. Student Interpretations of the Importance of Classroom Writing Assessment Constructs to their Performance on Classroom Assessments

Variable	*Very Important	*Important	*Somewhat important	*Minimally Important	*Not Important at all	Mean	Standard Deviation
Effective Use of Supporting Details (n=132)	65.2	31.8	3.0	0	0	4.62	0.55
Effective Sentence Construction (n=132)	65.9	26.5	5.3	2.3	0	4.56	0.70
Use Mechanics with Minimal Errors (n=132)	56.8	33.3	6.8	3.0	0	4.44	0.75
Use Grammar with Minimal Errors (n=132)	53.0	37.1	6.8	3.0	0	4.40	0.75
Develop Coherent, Focused, Shaped Ideas (n=132)	47.0	41.7	10.6	0.8	0	4.34	0.69
Effective Development of Voice (n=132)	50.8	32.6	15.2	0.8	0.8	4.32	0.81
Employ Diction Purposefully (n=132)	40.9	46.2	10.6	1.5	0.8	4.25	0.76
Develop Literary Interpretations (n=132)	41.7	41.7	14.4	2.3	0	4.23	0.78

* Responses in percentages.

Elements of construct under-representation in the diploma exam were viewed as less important than explicit elements of the construct, but more important than elements of construct irrelevant variance. Students reported that ability to critically re-examine their first drafts, their ability to restructure their first drafts, and their ability to negotiate a multiple draft writing process was important to their ability to do well on their classroom assignments.

Table 6.5. Student Interpretations of the Importance of Process Related Elements of Classroom Assessment Constructs to their Performance on Classroom Assessments

Variable	*Very Important	*Important	*Somewhat important	*Minimally Important	*Not Important at all	Mean	Standard Deviation
Critically Re-examine Draft One (n=132)	37.9	32.6	18.9	8.3	2.3	3.95	1.05
Restructure First Draft (n=132)	19.7	34.1	32.6	9.8	3.8	3.56	1.03
Multiple-draft writing process effectively (n=132)	18.2	30.3	32.6	12.9	6.1	3.42	1.11
Engage in Multiple-draft process (n=132)	13.6	31.8	25.8	22.0	6.8	3.23	1.14

* Responses in percentages.

Student ratings of writing process skills further reinforce the perception that writing in English 30-1 is focused within the traditional paradigm rather than in process oriented approaches. While thirty-eight percent of students believed that their ability to critically re-examine a first draft was very important for doing well on their writing assignments only twenty percent believed that their ability to restructure a first draft was very important to performing well on their classroom assignments. This response pattern would suggest that the focus on critical appraisal of first drafts is most focused on surface features of writing (mechanics, grammar, diction, syntax) rather than on the deep structures and the idea development that shapes effective writing. Similarly, only eighteen percent of students believed that an ability to utilize an effective multi-draft writing process was very important in leading them to greater success on their classroom assignments.

The final cluster of responses relate to instances both where elements extraneous to the construct and where ignored elements essential to the construct were seen to influence student performance on classroom assessments. Students felt that their teachers were more interested in aspects of the construct that were under-represented rather than aspects of the construct that introduced irrelevant variance into the exam scores. This parallels the experiences of Brian and Heather who did not actively teach to the exam, yet whose instruction was guided by the exam's explicit articulation of its construct.

Table 6.6. Student Interpretations of the Importance of Extraneous Construct Elements to their Performance on Classroom Writing Assessments

Variable	*Very Important	*Important	*Somewhat important	*Minimally Important	*Not Important at all	Mean	Standard Deviation
Write Effectively under Pressure (n=132)	14.4	25.8	38.6	15.2	6.1	3.27	1.08
Generate Ideas Quickly (n=132)	9.1	26.5	43.9	14.4	6.1	3.18	0.99
Create First Draft with a Minimum of Errors (n=132)	4.5	17.4	37.1	27.3	13.6	2.72	1.05

* Responses in percentages.

Though student responses suggest that the ability to write effectively under pressure and the ability to generate ideas quickly are not as important for doing well on classroom assignments as ability to polish writing might be, a healthy thirty-five to forty percent of student respondents indicated that these skills were either very important or important in performing well on classroom assignments. This suggests that in their classrooms a percentage of writing assignments were being designed to approximate exam like writing conditions. This pattern would be consistent with Brian and Anne's classrooms, where former exam questions (and at times in-class writing contexts) approximated the diploma exam context.

Students' critical reflection on their teacher's writing assignments suggests that the current traditional paradigm is alive and well in their classrooms where the focus on form and correctness is primary. The exam seems to feature heavily in students' beliefs about the skills that their teachers value. While writing process does seem to receive some emphasis, it is markedly less important than the skills related to creating essentially correct writing.

Student's Interpretation of the Exam's Construct

Students were also asked indirectly to interpret the exam's construct. They were asked to describe (interview) or rate (survey) the skills needed to perform well on the diploma exam. Students demonstrated a mature understanding of the exam's construct, perhaps a reflection of their teacher's focus on preparing them for the exam.

Of the students interviewed three believe that the exam is measuring one's ability to write quickly while under pressure; six believe that the exam was measuring their ability to understand the literature (presented in class and new literature presented on the exam) and apply it to the question posed in the exam; and seven believe that the exam was measuring their ability to develop and organize their ideas. Not a single student believes that the exam is measuring their ability to generate texts that were personally relevant or meaningful to themselves as writers.

Students' survey responses reflect student interview responses though with some noticeable differences. A primary difference relates to the issue of writing under pressure. While only three of ten interviewed students felt that one's ability to write under pressure was an important skill the exam was measuring, an overwhelming eighty-nine percent of students surveyed believed that one's ability to write under pressure was very important to performing well on the diploma exam while only three percent felt that this skill was somewhat or minimally important. Similarly, ninety-two percent of students believed that one's ability to generate ideas quickly was either important or very important in contributing to one's success on the diploma exam. One's ability to generate error-free first drafts was seen as less essential to performing well on the diploma exam: only sixty-three percent of students felt that this skill was either important or very important for achieving a high score on the exam. Students' focus on word or sentence level editing of a first draft might likely account for this drop in perceived importance.

Table 6.7. Student Interpretations of the Importance of Extraneous Construct Elements to their Performance on the English 30-1 Diploma Exam

Variable	*Very Important	*Important	*Somewhat important	*Minimally Important	*Not Important at all	Mean	Standard Deviation
Write Effectively under Pressure (n=132)	88.6	9.1	1.5	0.8	0	4.86	0.44
Generate Ideas Quickly (n=132)	73.5	18.2	6.1	1.5	0.8	4.62	0.74
Create First Draft with a Minimum of Errors (n=132)	29.5	33.3	18.2	8.3	10.6	3.63	1.28

* Responses in percentages.

Students also rated skills related to the construct quite highly. Over ninety percent of students rated all construct related skills (other than creation of voice) as being either important or very important to achieving a high mark on the diploma exam.

Table 6.8. Student Interpretations of the Importance of English 30-1 Diploma Exam Construct Elements for their Performance on the Diploma Exam

Variable	*Very Important	*Important	*Somewhat important	*Minimally Important	*Not Important at all	Mean	Standard Deviation
Effective Sentence Construction (n=132)	78.0	16.7	5.3	0	0	4.73	0.55
Effective Use of Supporting Details (n=132)	75	22	3	0	0	4.72	0.51
Use Mechanics with Minimal Errors (n=132)	74.2	20.5	3.8	1.5	0	4.67	0.62
Use Grammar with Minimal Errors (n=132)	68.9	25.8	3.8	1.5	0	4.62	0.64
Develop Coherent, Focused, Shaped Ideas (n=132)	64.4	30.3	5.3	0	0	4.59	0.59
Develop Literary Interpretations (n=132)	59.1	34.1	5.3	1.5	0	5.51	0.67
Employ Diction Purposefully (n=132)	62.1	28.8	7.6	0.8	0.8	4.51	0.74
Effective Development of Voice (n=132)	50.0	26.5	11.4	6.1	6.1	4.08	1.28

* Responses in percentages.

Elements of the construct not measured by the exam, on the other hand, were seen as being significantly less important to effective performance on the exam than all the skills students were asked to rate. As with the classroom assignments, one's ability to critically re-examine a first draft was seen to be more important than one's ability to restructure a first draft or to engage in a multiple-draft writing process. Forty-three percent of students believed that this skill led to better achievement on the diploma exam while only nine percent believed that one's ability to utilize a multiple draft writing process was very important for achieving a high score on the diploma exam.

Table 6.9. Student interpretations of the Importance of Process Related Elements of the Diploma Exam Construct to their Performance on the Diploma Exam

Variable	*Very Important	*Important	*Somewhat important	*Minimally Important	*Not Important at all	Mean	Standard Deviation
Critically Re-examine Draft One (n=132)	43.2	28.8	15.2	6.8	6.1	3.96	1.18
Restructure First Draft (n=132)	13.6	31.1	28.0	18.2	9.1	3.22	1.17
Multiple-draft writing process effectively (n=132)	9.1	22.0	25.8	26.5	16.7	2.80	1.21
Engage in Multiple-draft process (n=132)	7.6	19.7	19.7	30.3	22.7	2.59	1.25

* Responses in Percentages

Comparing student responses to their beliefs regarding skills needed to perform well on classroom assignments and skills required to perform well on the diploma exam reveals a perception of students that the two sets of assessments are measuring in some regards significantly different constructs. Most markedly, the ability to write under pressure was viewed by ninety-eight percent of students as being either important or very important for performing well on the diploma exam while only forty percent of students felt the same about this skill in relation to classroom assignments (a staggering difference of fifty-eight percent of students). Similarly, one's ability to effectively manage a multiple-stage writing process was seen to be important or very important for performance on the diploma exam (compared to classroom assignments) by eighteen percent fewer students. With the exception of one's ability to create a strong voice, the explicitly articulated elements of the construct were consistently rated as being either important or very important by a relatively equal number of students on both sets of assessment contexts, with the percentage of students rating these skills as important on the diploma exam only slightly higher (3% of respondents).

Clearly, students understand the exam construct quite well. They know the rubric well, for they clearly understand what skills it is measuring, but they also understand the unstated elements of the construct. They understand that the exam is measuring one's ability to write under pressure and not one's ability to utilize a multiple draft writing process. They understand too that the exam is measuring one's ability to generate relatively error free writing in that pressure writing context.

Writing Processes for Classroom Assignments

Generally speaking, the writing process described by the students interviewed for this study parallel the process envisioned in the current traditional paradigm. Eight of the ten students interviewed described using the following general process in their writing:

1. Generate material and/or ideas
2. Define a thesis
3. Write a draft
4. Engage in word or sentence level editing.

In terms of editing, however, students were split: half of the students claimed that once the draft was finished they did not like to revise or edit it while the other half of the students would revise to ensure ideas were clear or would revise if they felt that their drafting did not go smoothly. The following transcript of an interview with one of Heather's students exemplifies the process outlined above:

***James**: One thing that really helped [get me started on the assignment] was the first thing that was kind of mandatory. It helped give us a kick in the pants. We had to do research on stuff from our childhood. And that got ideas flowing and the nostalgia got my mind thinking. That was fun....As I was doing that I was writing down the ideas so I could use them later. After I had all those listed down I started – starting was the hardest part— once I thought of a line (we had to have an anchor line which is the main idea), I started with that and then I started throwing in the ideas that I had listed, and as I expressed them I filled up the page and then the ideas flowed and the entire thing came out once I was looking at the idea I had written down.*

***DHS**: Ok, did the ideas come while you were writing or before you did the writing?*

***James**: While I did the writing. When I first start I have no idea what I am talking about.*

***DHS**: So you develop the idea while you do the writing. Once you've got that first draft, or that first go-through done, do you go back through it at all or is it pretty much finished?*

***James**: I pretty much go through it to make sure I didn't do any spelling or grammar errors. But usually when I am writing I don't like to change my ideas because I am in a completely different mind set than when I was writing it, because my mind is completely different about five minutes after*

> *I completed writing it. So I am thinking I just will go through it, I don't want to edit it too much because then it usually ends up sounding like my ideas weren't flowing as well; so I will just make sure it is grammatically correct.*
>
> ***DHS**: Would you say that that process is similar or different from the process you use for essays?*
>
> ***James**: I use the same process.*

The processes described by James and other students reflect the research describing weak or developing writers done by Rose (1980/1994) and Sommers (1980/1994). These writers utilize a process that is rigidly defined, and guided by a dependence on certainty. Their processes are also modeled largely after the speech process and seem to view drafting as a means of expressing ideas in completed form. The students who engaged in more recursive processes, on the other hand, seemed to possess a greater flexibility in terms of deciding which process to use; while both describe a process that often reflects the process described by James, when the need arises they feel comfortable choosing a different process. Anne's strong writer describes his process as follows:

> ***DHS:** Could you describe the process that you use when you're writing an essay or when you're writing a piece of creative writing? What process do you use? What steps do you take?*
>
> ***John:** When I'm writing an essay I will sometimes make an outline of the basic points that I want to say, sometimes. Other times I just like to sit*

down and get the job done. I don't know if that's what most people do, or if that's the right way to do something, but it's the way that I end up doing most of the time. I think either way works as long as you review your work afterwards and look at it and say, "Is this is what I wanted to say?"

DHS: *And say you follow either approach and you come to the conclusion that this isn't necessarily what I wanted to say. What do you do from there?*

John: *Then I go back and change it. And I'd probably have to make out a bigger plan of why didn't I say this or how can this be changed to make it better.*

DHS: *Why have you chosen that freewriting approach as one you follow?*

John: *I guess that comes over the years doing homework tasks. I don't want to sit there all night and work out something. I'd just rather start at something, finish it, see if it's good, and if it's not then I'll change it to make it good. I don't do that all the time. But it's worked so far. I don't know if it'll work later on in life but if it doesn't then I'll have to change it.*

DHS: *So the other approach, the one that involves pre-planning and organizing, how well does that work for you?*

John: *That works fine. Just if I do it, I mean it's good. I think a lot of people use it and it really gets the structure down and I think it really helps them. I think I might be one of the outcasts that does it the other*

> *way but I mean if I were to use that structured thing, I think it would come out just as good. And a lot of people find that easier, much easier.*

The dialogue with John demonstrates his flexibility in terms of choosing his process while also demonstrating his allegiance to the idea that drafting is a process of putting pre-developed ideas to paper. For him, changing drafting process merely serves as a tool for better putting developed ideas on paper. The purpose of revision for John is to ensure that he got it right, that the ideas expressed on paper are the ideas he had intended to place there, rather than the ideas that developed as he wrote.

One of Heather's average writers, however, was the only interviewed student who described her process as being one of discovery. She describes a freewriting process in which she develops ideas, followed by a rereading and revision in which she searches for a means to organize and link those ideas. She describes her process as follows.

> ***DHS:*** *Can you take me through the process you use when you're writing an essay?*
>
> ***Kira:*** *The way I write them? I'll sit at the computer and I can try to start writing from the top and go to the bottom but I can't do it. Usually what happens is I'll start writing a jumbled mess and it will come to me and I'll get my thesis and I'll go, "Ok, I need to make use of this." And then I'll think of something for the introduction and then I'll think of something I can throw in the conclusion. So I'll get the important parts to me, the parts that I think...like they'll just come to me when I'm writing. Usually I don't*

even know, I'll just write something in the jumbled mess. And then I'll find it and I'll be, "Ok, this is awesome." And then I'll fill in the blanks. It's easy for me. But as far as sitting down and writing it all down from start to finish, I can't do that.

DHS: *So you start with a free writing kind of idea, and get ideas from there, your beginning, your end. How do you fill in the middle?*

Kira: *I just look at it and see what would relate more. Like usually I have a general idea so I'm not completely in the dark when I start it. I have kind of an idea, I just don't know how to wrap those thoughts around or what to throw in and make it make sense out of all of that.*

DHS: *So is it kind of a free write or is it more structured?*

Kira: *It's definitely more of a free write. And really the essays I tend to go over it a couple more times just to make sure.*

DHS: *So when you go over them, what are you looking at when you over them?*

Kira: *To make sure that...one of my problems that I used to have really badly, and it was really hard to get out of was I'd start writing and I'd get distracted. I'd start writing on other stuff and then when you come down to the end it's a bunch of all these unfinished ideas. So you're sitting there at the end going...what? And so I really had to work to change that. So a lot of times I'll write and then I'll leave it for like fifteen minutes and I'll come back and I'll read it again to make sure that it all makes sense. That*

> *is has an end point, that it has a general point, that it comes together. That there are facts and proof and all that.*

Kira's process utilizes writing as a means of discovery. She generates ideas through writing before being concerned with their organization and structure. Once she has developed her ideas, she searches for a means to organize them. Of the ten students interviewed, however, she is the anomaly.

The students who completed the survey too described a process that was less process-oriented and more traditionally based. Similar to the students who participated in the interviews, these students (eighty-six percent) believed that drafting was about putting preconceived ideas on paper in as clear a form as possible in the first draft. Sixty-one percent believed that clarity of expression was also an important goal of first draft writing. These beliefs help explain some of the behavior suggested by survey respondents: One third began their classroom writing assignments the day before they were due while one half completed their assignments on the day before they were due. This habit certainly does not lend itself to a robust or multi-draft writing process. In fact, forty-six percent of student respondents never or seldom wrote their assignments using a multi-draft process, while only seven percent always wrote using a multi-draft process. Given their beliefs about writing, student neglect of a multi-draft writing process should not be surprising, nor should it be surprising that forty-seven percent of respondents often or always attempted to complete their first drafts in as polished a form as possible, their goal was similar to that of many of my interview participants: to complete a first draft that did not require any revisions.

Table 6.10. Student Writing Habits

Variable	*Always	*Often	*Some-times	*Seldom	*Never	Mean	Standard Deviation
Complete assignments day before they are due (n=132)	12.9	39.4	31.1	11.4	5.3	3.43	1.03
Compete polished first drafts (n=132)	16.7	31.1	18.2	22.7	11.4	3.19	1.23
Begin Assignment Day before they are due (n=132)	7.6	25.8	30.3	25.8	10.6	2.94	1.12
Complete assignments using a multiple draft process (n=132)	7.6	18.2	28.0	34.1	12.1	2.75	1.12

* Responses in Percentages

Diploma Exam Writing Process

When it came to writing their diploma exam (or diploma preparation assignments) the students interviewed described their process as being similar to the process that they followed when writing their other assignments. What was interesting, however, was that half of these students commented on the issue of time constraints and its impact on process. They felt that the time constraints would either limit their ability to develop their ideas and/or their plan for their writing, or it would limit their ability to edit their papers. Six of the ten students believed that the exam context—time constraints and high pressure environment—reduced the quality of what they were able to produce. One of Brian's students reflects on the exam context as follows:

> ***DHS:*** *How well do you handle pressure, writing under pressure?*
>
> ***Tom:*** *Usually I try not to concern myself with the time because it makes you want to write quicker—like for my test when I was looking at the time and I was rushing so when I'm rushing too fast then I'll just wait a minute*

to actually realize how long a minute is. For me, writing under pressure, I'm not very good at it. I think I could be better if I did more of it. That's one thing I could work on, skill practice. You don't practice a lot in school and I think the more you practice it, the more you would be able to get better at it.

DHS: *How was yesterday for you in terms of pressure?*

Tom: *I got it done just in time but the thing is, it's not even the writing that is really the hardest. It's putting your ideas together within that time. Like writing your ideas out isn't the hardest part, it's deciding what you're going to write about because it's such an open sort of question.*

DHS: *So it's the thinking, that's the hard part, not the writing so much.*

Tom: *Yes, putting your ideas together.*

DHS: *How do you think your ability to handle pressure or not is going to affect your exam performance?*

Tom: *I think it will probably have a big impact. Well for me, like 5% is a lot, or 4%, in that range, it's a lot. So I think that because the amount of time I had, you don't have as much time to edit. You have a lot of thinking, less writing. So I think overall if we had more time you could formulate a lot better essay.*

Tom is Brian's strong writer. His experience draws to mind Berkenkotter's (1983/1994) study of Donald Murray's writing process. When placed in an exam context, this prolific writer was unable to produce more than three lines of text. If this high-pressure, time constrained process poses problems for this experienced writer, it should be no surprise that this same context causes significant problems for student writers as well. Clearly the issue of writing under pressure emerges both from the student interviews and from the survey results as being the primary issue students consider when thinking of this exam.

What is the Impact?

It is difficult to attribute student writing habits directly to the impact of this exam. Students' inability or unwillingness to engage in a significant revision or in writing as a means of discovery has been described in Sommers' (1980/1994) study comparing student writers with experienced adult writers. The problem, then, with the diploma exam is not necessarily that it actively teaches students to adopt poor understandings of what it means to write well, for they may, in fact, already possess these poor understandings, the problem is that the exam reinforces these poor understandings: By measuring student writing through a process that focuses on one's ability to create first drafts that are essentially correct and polished, the exam reinforces for students that this is how writing is done.

The exam does not challenge teachers or students to develop a more realistic understanding of what it means to write well.

CHAPTER 7: CONCLUSIONS AND IMPLICATIONS

The art of writing is the art of applying the
seat of the pants to the seat of the chair.
-Mary Heaton Vosse

Most of us can read the writing on the wall;
we just assume it's addressed to someone else."
-Ivern Ball

At the heart of this study rests perhaps the greatest irony in our era of accountability-driven education: Across North America, in an effort to improve the quality of education, governments have enacted test-based accountability programs but they have shown limited interest in determining whether these programs are in fact helping to improve or degrade educational quality. It seems that for the most part test-based accountability programs are built upon a blind faith in the integrity and validity of the assessment tools being used. Brennan (2006) in his introduction to *Educational Measurement* observes, "It certainly appears that a testing revolution is underway in this country [USA] that is based on the nearly unchallenged belief, with very little supporting evidence, that high-stakes testing can and will lead to improved education" (p. 9). His observation applies to the Canadian educational context as well. Many educators (Hillocks, 2002; Robinson, 2000), however, would argue that such faith is misplaced.

In this regard, I believe, the assessment profession has let down governments, educators, students, and the public: As a discipline, it has not actively pursued an agenda of accountability focused on the consequences stemming from the use of high-stakes tests; rather, its members have mostly quibbled over who should take responsibility for this work while avoiding the work itself. Brennan (2006), reflecting on the state of research in the field of assessment, observes that "validity theory is rich, but the practice of validation is often impoverished" (p. 8). He argues that test developers should focus their research on documenting the intended consequences of assessment use. He further argues that this research needs to be made available in a timely fashion. He concludes however,

> Unfortunately, publicly available, timely documentation related to validity arguments is often the exception rather than the rule. The uncomfortable reality is that if such documentation is clear, complete, and forthright, it will not always fully support validity arguments. (Brennan, 2006, p. 8)

In making this observation, Brennan taps into the heart of the issue: limited research regarding test validity and test use consequences stem from a fear of the answers such research might provide. Happily for some, questions unasked remain questions unanswered.

Questions Asked

This study, however, works against this grain. It builds upon a rich critical literacy foundation in the fields of curriculum and composition studies. It asks the questions too often ignored:

- Could a high-stakes writing assessment be flawed?
- If so, what impact might those flaws exert on teaching and learning?

- What contextual factors might enhance or mitigate an exam's impact on teaching and learning?

Questions Answered

Answers to the above questions lead to the following seven conclusions and eight recommendations:

A Limited Theory of Writing

- **The writing component of Alberta's English 30-1 diploma exam measures student writing on the basis of a flawed understanding of what it means to write well.**

Alberta's English 30-1 diploma exam assesses student writing on the basis of important qualities related to polished writing—idea development, supporting evidence, organization, voice, and clarity of expression—but it fails to adequately consider the process of writing. It provides students with a limited context in which to write—high pressure, tight time constraints, limited opportunities to plan, reconceptualize, and polish writing. Therefore, the exam assesses student writing using criteria appropriate for polished writing while asking students to complete their writing in a context that does not provide opportunities for the meaningful polishing of texts. The skill-set required to create polished first draft writing is very different from the skill-set required to polish writing through a robust, recursive process.

Negative Impact on Composition Pedagogy

- **The writing component of Alberta's English 30-1 diploma exam is having a negative impact on composition pedagogy in the contexts observed.**

Each of the three teachers observed developed and utilized writing assessments which reflected the construct being measured by the diploma exam: their assignments were literature-based, product-oriented, and (while not nearly as time restricted) focused on similar process skills being measured by the exam. As well, of the 132 students who participated in the survey component of this study, more than half reported receiving writing instruction often or always within the context of practice diploma exam questions and with a focus on formulaic organizational structures. Over 80% of these students identified core elements of the exam construct as being core elements of their classroom assessment constructs. Slightly fewer than 50% of these respondents' believed that an effective writing process would help them improve their performance on classroom assessments while 40% believed that learning to write under-pressure would help them improve their classroom assessment performances. These responses certainly suggest that the pedagogical conditions observed in the case-study settings were occurring in other classrooms as well.

Poor Reporting Practices

- **Alberta Education's method of reporting exam scores invites unwarranted comparisons between school-awarded marks and exam scores.**

By its own admission (McEwen, 1995, 2006) Alberta Education's diploma exam in English language arts education is consistently measuring a limited portion of the *Senior High ELA Program of Studies* (50% of the language arts, and 64% of total outcomes) yet Alberta Education publicly reports a school's average school report card scores alongside its average exam scores. This method of reporting invites comparisons between school and exam scores and it encourages media and other readers to make comparisons between these two sets of marks as well. For example, following the government's lead, the Fraser Institute has taken the government's reporting structure and has included comparisons between school and exam scores as an important factor in the way it ranks schools (Cowley & Easton, 2002). These comparisons, whether overt or implied, are unwarranted: The exam construct and the school report card constructs are significantly different. Because the diploma exam is only measuring a limited portion of the curriculum one should not expect school scores (based on a much more broad construct) to align with these exam scores. The current reporting structure used by Alberta Education, however, suggests the opposite.

External Pressure Drives Teaching to the Test

- **Increased pressure to align exam scores and school scores leads to an increase in teachers' focus on teaching and assessing (in the classroom context) the skill-set captured by the exam's construct.**

Teachers reported varying degrees of both overt and subtle pressures from their school boards and their school administrators to ensure that their school awarded marks and their exam scores align with one another: As this pressure increases so does the teachers' focus on teaching to the exam's construct. Teachers recognize that the best (and perhaps only) way to ensure alignment of scores is to ensure that the constructs being measured by the exam and by classroom assessments are in fact similar constructs.

Critical Awareness Influences Pedagogy

- **Critical awareness of the exam's flaws lead to more overt forms of test preparation.**

The teachers involved in this study exhibited a range in terms of their critical perspectives on the diploma exam's construct. The teacher who had received the most advanced instruction in methods for teaching writing was the most critical of the exam's construct. The teachers whose education in composition pedagogy was more limited were less concerned about the exam's construct. As critical awareness of the exam's construct increased so did the act of overtly teaching to the test. (This relates to the previous conclusion in that pressure to align school marks and exam scores led teachers who were critical of the exam to conclude the best way to ensure alignment of scores was to measure in school the same construct the exam was measuring—this understanding intensifies as critical awareness of the exam increases.) The teachers who were less aware of the exam's construct flaws exhibited fewer overt acts of teaching to the exam yet their pedagogy more naturally aligned with the exam's construct.

Inflation of Test Scores not a Primary Motivation

- **Teaching to the test was not driven by a desire to increase test scores.**

In the assessment literature (Koretz & Hamilton, 2006) teachers are frequently criticized for the practice of teaching to the test in an effort to artificially inflate test results. In the cases observed where teaching to the test occurred, the desire to increase test scores was not a motivating factor. Two motivating factors were identified: First, as discussed earlier, teaching and assessment focused on the exam construct stemmed from a desire to ensure that average school scores and exam scores aligned with one another. Second, teachers felt it was their professional responsibility to ensure that their students were well prepared for a high-stakes exam that carried the potential to significantly effect the lives of students (university applications depended to varying degrees on exam performance). In an effort to ensure their students were prepared for this exam, teachers either overtly or implicitly prepared their students for the skill-set being measured by the exam.

Diploma Exam Supports Limited Learning

- **The writing construct of the writing component of the English 30-1 diploma exam supports limited student learning.**

The data from the case studies and the surveys are clear. The vast majority of student participants in this research utilized a problematic and limited writing process: they generated ideas, drafted materials, and engaged in surface level editing. This process conforms with the process they believe will lead to success on the diploma exam. Interestingly, 75 % of the students' who responded to the survey believed that an effective writing process was a foundational element of good writing. The effect of the exam can be seen in the fact that in spite of this belief, only 25% of students either often or always engaged in a multiple-draft process when writing. While other factors (such as tendency to procrastinate) certainly would contribute to students' choice of process, it is clear that the process that students are utilizing in their writing is being supported by the exam.

Implications

Driving or Supporting?

Given that the stated purpose of test-based accountability is to improve the quality of education, at the very least one should expect that high-stakes tests implemented within an accountability-framework would support excellence in teaching and learning. The writing portion of the English 30-1 diploma exam seems to not be accomplishing this goal. The exam's flawed construct supports a perspective on writing process which emerges from the thinking behind the Current-Traditional-Paradigm, a way of thinking about writing that was largely discredited by the 1980's. In the case of Anne—whose teacher education program was embedded in a process-oriented approach—it is clear that her experience with the diploma exam has driven her toward adopting a pedagogy which stems from the Current-Traditional-Paradigm. In the case of Brian and Heather—teachers who had limited education in process-oriented approaches to teaching writing—the exam supports their pedagogical focus, one largely built around the Current-Traditional-Paradigm. In Brian's and Heather's cases the exam clearly does not encourage them to take a more progressive approach to teaching writing. In Anne's case, the exam's construct can be said to be driving a limited form of composition pedagogy; in the other cases, the exam's construct can be said to be supporting a limited form of pedagogy. In all cases the consequences for students are not good.

Just as the exam supports poor pedagogy, so too can it be said to support poor writing habits and skill development in students. The limited process used by the vast majority of this study's student participants is one which often leads to limited idea development and frequent frustrations with writing. An exam which encourages process-oriented methods of writing would prompt students and teachers to take a more progressive approach to writing.

Necessity of Redesign of the Diploma Exam

- **The writing component of Alberta's English 30-1 diploma exam is in need of significant redesign.**

Sheppard (2006) eloquently argues for a philosophical alignment of pedagogy and assessment. The evidence in my study points clearly to a failure of alignment, and points to significant consequences stemming from this failure. While compositionists have made significant steps forward in terms of composition pedagogy in the past forty years, Alberta's English 30-1 diploma exam is largely based upon an old-paradigm model of writing. If the goal of test-based accountability is to improve teaching and learning Alberta Education needs to redesign its diploma exam so that the exam expects robust processes, encourages risk-taking and promotes exploration.

Method of Score Reporting Requires Revision

- **Alberta Education's current method of reporting exam scores needs to be modified to avoid unwarranted comparisons between school awarded marks and diploma exam scores.**

As indicated above, the current reporting structure used by Alberta Education invites problematic comparisons of exam scores and school marks. A new reporting structure which makes explicitly clear that these comparisons are not warranted but rather that they are unhelpful is necessary. Additionally, school boards and school administrators who implement policies requiring the alignment of school awarded marks and diploma exam scores should revise this policy.

More Validity-based Research Required

- **Scholars in the field of assessment need to take up their responsibilities regarding validity-based research.**

The debate regarding who should take responsibility for research into the validity and consequences of assessment presents a false dilemma. Edward Said (1996) argues that the academic, as public intellectual, has a responsibility to work for the public good. Specialists in the field of assessment have the knowledge and the ability to implement significant research programs which explore issues regarding the consequences of assessment on systems of education. Certainly in today's educational climate, funding agencies such as the Social Sciences and Humanities Research Council could likely provide funding for this research. Without an active body of validity research being conducted we will continue to implement a test-based accountability model of education without ever knowing if tests are serving to improve or degrade educational quality.

Critical Assessment Literacy Required

- **Teacher education programs should ensure that teachers have the skills needed to engage critically with assessments.**

Much of the work in classroom assessment focuses on helping teachers develop the skills needed to develop sound assessments in their own classrooms. While this skill-set is important, pre-service and in-service teachers need to be trained more explicitly in issues and implications related to validity theory. Knowledge of the ethical implications of validity theory will enable teachers to better advocate for assessments which reflect the teaching and learning goals articulated in government directed curricula.

Research Implications

National Comparative Study Needed

- **A national comparative study of composition pedagogy in Grade 12 English language arts classrooms in each province would strengthen our understanding of the effects of high-stakes writing assessment on teaching and learning.**

This study suggests that Alberta's English 30-1 diploma exam supports a limited form of composition pedagogy. A broader view, however, would add strength to this position.

Canada is unique in terms of its education system. In western and eastern Canada two related common curriculum frameworks have been developed which establish (within their jurisdictions) the common elements for English language arts curricula across their respective regions. For example, under the Western and Northern Canadian Protocol, British Columbia, Alberta, Saskatchewan, Manitoba, The North West Territories, Yukon, and Nunavut have developed common ELA curricula based upon the Northern and Western Canadian Protocol's Common Curriculum Framework for English Language Arts. Each province and territory, however, has maintained control of its own assessment or accountability programs so that British Columbia, Alberta, Manitoba, and Saskatchewan each have very different standardized writing assessments at the grade 12 level. Manitoba, for example, has a writing exam that is built upon a linear process model with students completing this exam over the course of five days. A study across these jurisdictions would be valuable: If in each province the pedagogy and student learning is the same, we can attribute this to common curricula and common perspectives on pedagogy; however if in each province the pedagogy is different, reflecting the differences in each province's standardized writing assessments, these findings would warrant a stronger argument regarding the impact of standardized writing assessment on teaching and learning in Canada.

Research Across all Grade Levels is Necessary

- **A study of writing instruction and standardized assessment from Grade 2 to Grade 12 in Alberta would help us understand the impact of Alberta's standardized writing assessment program on our elementary and secondary system as a whole.**

Alberta Education has implemented standardized writing assessments at the grade 3, 6, and 9 levels. Each of these exams is built upon a model of writing that is similar to the grade 12 diploma exam's model. A study similar to this one but at the grade 3, 6, and 9, and 30-2 (grade 12 high school completion English) levels will help create a more complete picture of this assessment program's impact on writing instruction in elementary and secondary schools in Alberta.

Renewed Research into Writing Assessment Design

- **Renewed research into writing assessment design.**

Improvements or reforms in writing assessment were driven largely by research into new and innovative assessment designs. More recent strides forward toward the use of portfolio assessments have largely been reversed in recent years. Psychometric concerns around the costs and the reliability of portfolio assessments have largely precipitated this shift. Research into new designs for portfolio assessments is needed as we search for portfolio assessment designs which enhance the reliability of these assessments without sacrificing their validity.

Limitations and Promises

The educational measurement community has struggled for some time now with the question of test use consequences. This study, I believe, will help move that discussion forward: It demonstrates how questions of test use consequences can be effectively explored; it demonstrates the importance of conducting assessment oriented research within a qualitative framework; and it points to the consequences stemming from failures of current test-based models of educational accountability.

Its relative solitude is perhaps one of the most significant limitations of this study. A broader range of studies within the Alberta context, within the Canadian context, and across international contexts is needed to enhance the potential impact of research into the consequences of standardized assessments on teaching and learning. Future research designed to examine similar questions with a similar focus on participant contexts would help to support the findings of this study.

A related limitation of this study, enhanced by its relationship to the field of educational measurement, is the relatively limited sample size. Large-scale quazi-experimental research is generally speaking the privileged method of research in the field of educational measurement. While I plan to continue to frame my research within a qualitative framework I hope to conduct future mixed-methods studies on a much larger scale.

In spite of these limitations, this study does hold out promise to teachers and students: the promise that their frustrations regarding the limitations and failures of test-based accountability systems are being explored with a view to promoting the improvement of these systems. This study suggests to teachers and students that standardized assessments (and those that develop and implement them) too can be held accountable for their impacts on educational quality. This promise, however, remains a distant one. Studies of test use consequences are rare; individual studies may point to future promise but they require a larger body of work both to support them and to enhance their potential to shape future educational landscapes.

BIBLIOGRAPHY

AERA & NCMUE (1955). *Technical recommendations for achievement tests.* Washington: AERA.

Alberta Education (2003). *English language arts (senior high).* Retrieved February 19, 2007 from http://www.education.gov.ab.ca/k%5F12/curriculum/bySubject/english/srhelapofs.pdf

Alberta Education (2003). *Senior high school English language arts guide to implementation.* Retrieved February 14, 2007 from http://www.education.gov.ab.ca/k%5F12/curriculum/bySubject/english/

Alberta Education (2004). *Diploma examinations program.* Retrieved March 29, 2004 from http://www.learning.gov.ab.ca/k_12/testing/diploma/dib_gib/examinationprogram.asp

Alberta Education (2005). *English language arts 30-1 June 2005 writing assignments.* Retrieved December 15, 2006 from http://www.education.gov.ab.ca/k%5F12/testing/diploma/bulletins/examples_stand/default.asp

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Anderson, B. (2006, May). *Does every child count? British Columbia's educational accountability in 2006.* Paper presented at the annual meeting of Canadian Society for Studies in Education, Toronto, ON.

Atwood, M. (2002). *Negotiating with the dead: A writer on writing.* New York: Cambridge University Press.

Bartholomae, D. (1995). Writing with teachers: A conversation with Peter Elbow. [Electronic Version] *College Composition and Communication, 46*(1) 62-71.

Berkenkotter, C. (1994). Decisions and revisions: The planning strategies of a publishing writer. In S. Perl (Ed.), *Landmark essays on writing process*. (pp. 107-112). Ann Arbor: Hermagoras Press. (Original work published 1983)

Bevington, D. (Ed.). (1980). *The complete works of Shakespeare* (3rd ed.). Glenview, IL: Scott, Foresman.

Brennan, A. L. (Ed). (2006) *Educational measurement* (4rth ed.). Westport, CT: Praeger Publishers.

Brenan, A. L. (2006). Perspectives on the evolution and future of educational measurement. In Robert L. Brennan (Ed.) *Educational measurement* (4rth ed., pp. 1-16). Westport, CT: Praeger Publishers.

Britton, J. N., Burgess, T., Martin, N., McLeod, A., Rosen, H. (1975). *The Development of writing abilities, 11-18.* London: McMillan.

Brown, R. (1978). What we know and how we could know more about writing ability in America. *Journal of Basic Writing, 1,* 1-6.

Calkins, L. M. (1986). *The art of teaching writing.* Portsmouth, NH: Heinemann.

Clandinin, D. J. & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research.* San Francisco: Jossey-Bass Publishers.

Coe, R. M. (1990). *Process, form, and substance: A rhetoric for advanced writers.* Englewood Cliffs, NJ.: Prentice Hall.

Coles, W. E. J. & Vopat, J. (1985). *What makes writing good: A multiperspective.* Toronto: D. C. Heath and Company.

Cook, T. D. & Campbell, D. T. (1979*). Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.

Couture, B. (1999). Modeling and emulating: Rethinking agency in the writing process. In T. Kent's (Ed.) Post-*process theory: Beyond the writing process paradigm.* (pp. 30-48). Carbondale IL: Southern Illinois University Press.

Cowley, P. & Easton, S. (2002). *Report card on Alberta's high school's.* Vancouver: The Fraser Institute. Retrieved on February 21, 2007 from, http://www.fraserinstitute.ca/admin/books/chapterfiles/Introduction%20and%20notes-1ABSRIntro.pdf.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches*. (2nd ed.) London: Sage publications.

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer and H. Braun (Eds.), *Test validity* (pp. 3-17). Hillsdale, NJ: Erlbaum.

Cronbach, L. J. (1989). Construct validation after 30 years. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 147-171). Urbana: University of Illinois Press.

Crundwell, R. M. (2005). Alternative strategies for large scale student assessment in Canada: Is value added assessment one possible answer? Canadian Journal of Educational Administration and Policy, 41. Retrieved November, 2006, from http://www.umanitoba.ca/publications/cjeap/articles/crundwell.html

D' Angelo, F. J. (1975). *A Conceptual theory of rhetoric.* Cambridge, MA: Winthrop Publishing Co.

Denzin, N. K. & Lincoln, Y. S. (*Eds.*). (1998). *The landscape of qualitative research: Theories and issues*. Thousand Oaks CA: Sage.

Denzin, N. K. & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research. In Norman K. Denzin & Yvonna S. Lincoln (Eds.) *The Sage handbook of qualitative research.*(3rd ed., pp. 1-32). Thousand Oaks: Sage Publications.

Denzin, N. K. & Lincoln, Y. S. (Eds.). (2005). *The Sage handbook of qualitative Research.*(3rd ed.). Thousand Oaks: Sage Publications.

Dickens, C. (2000). *A Tale of Two Cities.* London: Penguin Books.

Didion, J. (1994). Why I write. In G. H. Muller & A. F. Crooks (Eds.), *Major modern essayists* (2nd ed., pp. 224-230). Englewood Cliffs: Prentice Hall.

Dillard, A. (1989). *The writing life*. New York: Harper and Row.

Downing, S. M. (1996) A model for evaluating high-stakes testing programs: Why the fox should not guard the chicken coop. *Educational Measurement: Issues and Practice, 15,* 5-12.

Earl, L. (1999). Assessment and accountability in education: Improvement or surveillance? *Education Canada, 39* (3), 4-6.

Elbow, P. (1973). *Writing without teachers*. New York: Oxford University Press.

Elbow, P. (1981). *Writing with power: Techniques for mastering the writing process*. New York: Oxford University Press.

Elbow, P. (1995). Being a writer vs. being an academic: A conflict in goals. [Electronic Version] *College Composition and Communication, 46*(1) 72-83.

Emig, J. (1971). *The composing processes of twelfth graders.* (NCTE research report No. 13) Urbana: NCTE.

Erdrich, L. (1982). *Love medicine.* New York: Holt, Rinehart, and Winston.

Ferlinghetti, L. (1968). *A Coney island of the mind.* New York: New Directions.

Fitzgerald, F. S. (2003). *The great Gatsby.* New York: Scribner Paperback Fiction.

Flinders, D. J. & Eisner, E. W. (1994). Educational criticism as a form of qualitative inquiry. *Research in the Teaching of English, 28*, 341-357.

Flowers, L. & Hayes, J. R. (1994). The cognition of discovery: defining a rhetorical problem. In S. Perl (Ed.), *Landmark essays on writing process*. (pp. 63-74). Ann Arbor: Hermagoras Press. (Original work published 1980)

Freedman, J. (1995). Comments on Meaghan and Casas' "On the testing of standards and standardized achievement testing." [Electronic Version] *Interchange, 26*(1) 59-71.

Gautron, H. (2006). *Accountability reporting in Quebec schools since 1995.* Paper presented at Canadian Society for Studies in Education, Toronto, ON.

Gordon, B., Engelhard, G. Jr., Gabrielson, S., & Bernknope, S. (1996). Conceptual issues in equating performance assessments: Lessons from writing assessment. *Journal of Research and Development in Education, 29,* 81-88.

Graves, D. (1994). An examination of the writing processes of seven-year-old children. In S. Perl (Ed.), *Landmark essays on writing process*. (pp. 23-38). Ann Arbor: Hermagoras Press. (Original work published in 1975)

Green, D. R. (1998). Consequential aspects of the validity of achievement tests: A publisher's point of view. *Educational Measurement: Issues and Practice, 17,* 16-19, 34.

Green, J. C. (2005). The generative potential of mixed methods inquiry. [Electronic Version] *International Journal of Research & Method in Education, 28*(2) 207-211.

Grumet, M. (1992). The language in the middle. [Electronic Version] *Liberal Education, 78*(3). 2-6.

Guba, E. G. & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In Norman K. Denzin & Yvonna S. Lincoln (Eds.) *Handbook of qualitative research.* (pp. 105-117). Thousand Oaks: Sage Publications.

Guba, E. G. & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. . In Norman K. Denzin & Yvonna S. Lincoln (Eds.) *The Sage handbook of qualitative research.* (3rd ed., pp. 191-216). Thousand Oaks: Sage Publications.

Hairiston, M. (1986). Different products, different processes: a theory about writing.

College Composition and Communication, 37, 442-452.

Hairiston, M. (1994). The winds of change: Thomas Kuhn and the revolution in the teaching of writing. In S. Perl (Ed.), *Landmark essays on writing process*. (pp. 113-126) Ann Arbor: Hermagoras Press. (Original work published 1982)

Hamp-Lyons, L. (2002). The scope of writing assessment. [Electronic Version] *Assessing Writing, 8,* 5-16.

Hillocks Jr., G. (2002). *The testing trap: How state writing assessments control learning.* New York: Teachers' College Press.

Hillocks Jr., G. (2003). Fighting back: Assessing the assessments. *English Journal, 92,* 63-70.

Hoffman, J. V., Paris, S. G., Salas, R., Patterson, E., & Assaf, L. (2003). High-stakes assessment in the language arts: the piper plays, the players dance, but who pays the price. In J. Flood, J. M. Jenson, D. Lapp, & J. Squire (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 619-630). Mahwah, N. J.: L. Earlbaum Associates.

Hoxby, C. M. (2002). *School choice and school productivity (or could school choice be a tide that lifts all boats?).* NBER Working Papers 8873, National Bureau of Economic Research, Inc. Online: http://ideas.repec.org/e/pho46.html .

Huot, B. (1990). Reliability, validity, and holistic scoring: What we know and what we need to know. *College Composition and Communication, 41,* 201-213.

Huot, B. (1996). Toward a new theory of writing assessment *College Composition and Communication 47*, 549-567.

Jaafar, S. B. (2006, May). *Accountability, then, now and next.* Paper presented at the annual meeting of Canadian Society for Studies in Education, Toronto, ON.

Johnson, B. & Christensen, L. (2004). *Educational research: Quantitative, qualitative and mixed approaches*. (2nd ed.) Toronto: Pearson.

Kane, M. T. (2006). Validation. In Robert L. Brennan (Ed.), *Educational measurement* (4rth ed., pp. 17-66). Westport, CT: Praeger Publishers.

Kent, T. (1999) (Ed.) Post-*process theory: Beyond the writing process paradigm.* Carbondale IL: Southern Illinois University Press.

Kessler, K. (2005). Composing for delivery. [Electronic Version] *English Journal 95*(2). 89-96.

King, S. (2000). *On writing: A memoir of the craft*. New York: Scribner.

Kinneavy, J. (1971). *A Theory of discourse.* Englewood Cliffs, NJ: Prentice Hall.

Koretz, D. M. & Hamilton, L. S. (2006). Testing for accountability in K-12. In Robert L. Brennan (Ed.), *Educational measurement* (4rth ed., pp. 531-578). Westport, CT: Praeger Publishers.

Lane, S., Parke, C. S., & Stone, C. A. (1998). A framework for evaluating the consequences of assessment programs. *Educational Measurement: Issues and Practice, 17,* 24-28.

Remarque, E. M. (1984). *All quiet on the western front*. New York: Ballantine Books.

Rex, L. A. & Nelson, M. C. (2004). How teachers' professional identities position high-stakes test preparation in their classrooms. [Electronic Version] *Teachers College Record, 106*(6), 1288-1331.

Lincoln, Y. S. (1995). Emerging criteria for quality in qualitative and interpretive research. *Qualitative Inquiry 1,* 3, 275-289.

Linn, R. L. (1998). Partitioning the responsibility for the evaluation of the consequences of assessment programs. *Educational Measurement: Issues and Practice, 17,* 28-30.

Linn, R. L. & Kiplinger, V. L. (1992). Cross-state comparability of judgements of student writing: results from the new standards project. [Electronic Version] *Applied Measurement in Education, 5,* 89-110.

Linn, R. L. (1997). Evaluating the validity of assessments: The consequences of use. *Educational Measurement: Issues and Practice, 16,* 14-16.

Mabry, L. (1999). Writing to the rubric. *Phi Delta Kappan, 80,* 673-680.

Madraso, J. (1993) Proofreading: The skill we've neglected to teach. *English Journal 82*(2) 32-42.

Maquire, T., Hattie, J., & Haig, B. (1994). Construct validity and achievement assessment. *The Alberta Journal of Educational Research, 40,* 109-126.

Mason, J. (2006). Mixing methods in a qualitatively driven way. *Qualitative Research 6*(1) 9-25.

Matsuda, P. K. (2003). Process and post-process: A discursive history. [Electronic Version] *Journal of Second Language Writing 12,* 65-83.

Maxwell, J. A. (1992). Understanding validity in qualitative research. [Electronic Version] *Harvard Educational Review 62*(3), 279-300.

McDowell, L. & Sambell, K. (1999). Fitness for purpose in the assessment of learning: students as stakeholders. [Electronic Version] *Quality in Higher Education, 5,* 107-123.

McEwen, N. (1995) Educational accountability in Alberta. *Canadian Journal of Education, 20,* 1, 27-44.

McEwen, N. (2006). *The impact of educational accountability in Alberta since 1995.* Paper presented at the annual meeting of Canadian Society for Studies in Education, Toronto, ON.

Meaghan, D. E., & Casas, F. R. (1995). On standardized achievement testing: Response to Freedman and Wilson and a last word. [Electronic Version] *Interchange, 26*(1), 81-96.

Mehrens, W. A. (1997). The consequences of consequential validity. *Educational Measurement: Issues and Practice, 16,* 16-18.

Merriam, S. B. (2001) *Qualitative research and case study applications in education.* San Francisco: Jossey-Bass Publishers.

Mertens, D. M. (2005) *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods.* (2nd Ed). Thousand Oaks: Sage Publications.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18,* 5-11.

Messick, S. (1990). *Validity of test interpretation and use* (Rep. No. ETS-RR-90-11). Princeton, NJ: Educational Testing Service.

Messick, S. (1994). *Alternative modes of assessment, uniform standards of validity* (Rep. No. ETS-RR-94-60). Princeton, NJ: Educational Testing Services.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23,* 13-23.

Messick, S. (1995). Standards of validity and validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14,* 5-8.

Messick, S. (1996). *Validity and washback in language testing* (Rep. No. ETS-RR-96-17). Princeton, NJ: Educational Testing Service.

Moffett, J. (1968). *Teaching the universe of discourse.* Boston: Houghton Mifflin.

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research, 62,* 229-258.

Moss, P. A. (1994). Can there be validity without reliability. *Educational Researcher, 23,* 5-12.

Moss, P. A. (1995). Themes and variations in validity theory. *Educational Measurement: Issues and Practice, 14,* 5-13.

Moss, P. A. (1998). The role of consequences in validity theory. *Educational Measurement: Issues and Practice, 17,* 6-12.

Murray, D. (1968). *A writer teaches writing*. Boston: Houghton Mifflin.

Murray, D. (1990). *Shoptalk: Learning to write with writers*. Portsmouth: Heineman.

Nunan, S. L. (2005). Forgiving ourselves and forging ahead: Teaching grammar in a new millennium. *English Journal 94*(4) 70-76

Orwell, G. (1994). Why I write. In G. H. Muller & A. F. Crooks (Eds.), *Major modern essayists* (2nd ed., pp. 90-97). Englewood Cliffs: Prentice Hall.

Palmer, J., Brestler, L., & Cooper, D. E. (2001). *Fifty major thinkers on education: from Confucius to Dewey.* New York: Routledge.

Palys, T. (1997). *Research decisions: Quantitative and qualitative perspectives.* (2nd ed.) Toronto: Harcourt Brace Canada.

Paraskevas, C. (2004). The craft of writing: Breaking Conventions. *English Journal 93*(4) 41-47.

Perl, S. (Ed.). (1994). *Landmark essays on writing process.* Ann Arbor: Hermagoras Press.

Perl, S. (1994). The composing processes of unskilled college writers. In S. Perl (Ed.), *Landmark essays on writing process.* (pp. 39-62). Ann Arbor: Hermagoras Press. (Original work published 1979)

Perry, L. (2006, May). *Accountability in Newfoundland and Labrador education: Can we maintain credibility?* Paper presented at the annual meeting of Canadian Society for Studies in Education, Toronto, ON.

Perry-Fagan, L. (1995). Performance accountability in the Newfoundland school system. *Canadian Journal of Education, 20,* 1, 65-76.

Petraglia, J. (1999). Is there life after process? The role of social scientism in a changing discipline. In T. Kent's (Ed.) *Post-process theory: Beyond the writing-process paradigm.* (pp. 49-64). Carbondale, IL: Southern Illinois University Press.

Popham, W. J. (1997). Consequential validity: Right concernBwrong concept. *Educational Measurement: Issues and Practice, 16,* 9-13.

Popham, W. J. (1999). Where large scale educational assessment is heading and why it shouldn't. *Educational Measurement: Issues and Practice, 18,* 13-17.

Principles of fair assessment practices for education in Canada. (1993). Edmonton, Joint Advisory Committee. Center for Research in Applied Measurement and Evaluation, University of Alberta. Retrieved on February, 21, 2007 from, http://www.education.ualberta.ca/educ/psych/crame/files/eng_prin.pdf

Punch, M. (1998) Politics and ethics in qualitative research. In N.K. Denzin & Y. S. Lincoln (Eds.) *The landscapes of qualitative research: Theories and issues* (pp. 156 - 184). Thousand Oaks: Sage Publications.

Purves, A. C. (1992). Reflections on research and assessment in written composition. *Research in the Teaching of English, 26,* 108-122.

Reckase, M. D. (1998). Consequential validity from the test developer's perspective. *Educational Measurement: Issues and Practice, 17*, 13-16.

Remarque, E. M. (1996). *All Quiet on the Western Front.* London: Vintage.

Resnick, L., & Resnick, D. 1992, Assessing the thinking curriculum: New tools for educational reform. In B.R.Gifford & M.C.O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 35-75) Boston: Kluwer Academic Publishers.

Robinson, S. (2000). "Miles to go. . .": Assessment and secondary English. In B. R. C. Barrell & R. F. Hammett (Eds.), *Advocating change: contemporary issues in* *subject English* (pp. 254-280). Toronto: Irwin Publishing.

Rose, M. (1994). Rigid rules, inflexible plans, and the stifling of language: A cognitivist analysis of writer's block. In S. Perl (Ed.), *Landmark essays on writing process.* (pp. 85-98). Ann Arbor: Hermagoras Press. (Original work published 1980)

Rosenblatt, L. M. (1938). *Literature as experience.* New York: D. Appleton Century.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work.* Carbondale, IL: Southern Illinois University Press.

Rosenblatt, L. M. (1993). The transactional theory: against dualisms. *College English, 55*(4), 377-386.

Rosenblatt, L. M. (1998). Readers, texts, authors. *Transactions of the Charles S. Peirce Society 34*(4), 885-921.

Russell, D. (1999). Activity theory and process approaches: Writing (power) in school and society. In T. Kent's (Ed.) *Post-process theory: Beyond the writing-process paradigm.* (pp. 80-95). Carbondale, IL: Southern Illinois University Press.

Rymer, J. (1988). Scientific composing processes: How eminent scientists write journal articles. *Advances in Writing Research, 2*, 211-250.

Said, E. W. (1996). *Representations of the intellectual.* New York: Vintage Books.

See, C. (2002). *Making a literary life*. New York: Random House.

Shafer, G. (2004) Reforming writing and rethinking correctness. *English Journal 94*(1) 66-72.

Shepard, L. A., (1997). The centrality of use and consequences for test validity. *Educational Measurement: Issues and Practice, 6,* 5-8.

Shepard, L. A., (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Shepard, L. A., (2006). Classroom assessment. In Robert L. Brennan (Ed.), *Educational measurement* (4rth ed., pp. 623-646). Westport, CT: Praeger Publishers.

Simmer, M. L. (2000). *A joint position statement by the Canadian Psychological Association and the Canadian Association of School Psychologists on the Canadian press coverage of the province-wide achievement test results.* Ottawa: Canadian Psychological Association. Online: http://www.cpa/documents/joint_position.html .

Slomp, D. (2005). Teaching and assessing language skills: Defining the knowledge that matters. *English Teaching Practice and Critique, 4*(3) 141-155. Retrieved February 14, 2007, from http://education.waikato.ac.nz/research/files/etpc /files/2005v4n3art8.pdf

Smith, M. L. & Fey, P., (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education, 51(*5), 334-344

Sommers, N. (1994). Revision strategies of student writers and experienced adult writers. In S. Perl (Ed.), *Landmark essays on writing process*. (pp. 75-84). Ann Arbor: Hermagoras Press. (Original work published 1980)

Stafford, K. (2003). *The muses among us: Eloquent listening and other pleasures of the writer's craft*. London: University of Georgia Press.

Stake, R. E. (2000). Case studies. In N. K. Denzin & Y. V. Lincoln (Eds.) *Handbook of qualitative research* (2nd Ed. pp. 435 - 454). Thousand Oaks CA: Sage

Stake, R. E. (2005). Qualitative case studies. . In Norman K. Denzin & Yvonna S. Lincoln (Eds.) *The Sage handbook of qualitative research.*(3rd ed., pp. 443-466). Thousand Oaks: Sage Publications.

Stake, R. E. (2006). *Multiple case study analysis.* New York: The Guilford Press.

Stiggins, R. J. (1999). Assessment , student confidence, and school success. *Phi Delta Kappan, 81*(3), 191-198.

Tashakkori, A. & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks CA: Sage.

Tashakkori, A. & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in the social and behavioral sciences*. Thousand Oaks CA: Sage.

Tobin, L. (1994). Introduction: How the writing process was born—and other conversion narratives. In L Tobin & T. Newkirk's (Eds.). *Taking stock: The writing process movement in the '90's.* (pp. 1- 16). Portsmouth NH: Boynton/Cook Publishers, Heinemann.

Tobin, L. (2001). Process pedagogy. In G. Tate, A. Rupiper, & K. Schick (Eds.). *A guide to composition pedagogies,* (pp. 1-18). New York: Oxford University Press, 2001.

Volante, L. (2006, May) *Assessment, accountability, and educational reform in Ontario.* Paper presented at the annual meeting of Canadian Society for Studies in Education, Toronto, ON.

Vygotsky, L. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge, MIT

Press.

Warner, A. L. (1993). If the shoe nolonger fits, wear it anyway? *English Journal 82*(5) 76-81.

Weigle-Cushing, S. (2002). *Assessing writing*. New York: Cambridge University Press.

Wiesel, E. (1960). *Night* (S. Rodway, Trans.). New York : Bantam Books.

Wiggins, G. (1994). The constant danger of sacrificing validity to reliability: Making writing assessment serve writers. *Assessing Writing, 1,* 129-139.

Wiley, D. E. (1991). Test validity and invalidity reconsidered. In R. E. Snow & D. E. Wiley (Eds.), *Improving inquiry in the social sciences: A volume in honor of Lee. J. Cronbach* (pp. 75-107). Hillsdale, NJ: Earlbaum.

Yackulic, R. A. & Noonan, B.W. (1999). *Measurement of the full scope of learning*. Presented at the Pan-Canadian Education Research Agenda Symposium, Ottawa: February 16-17, 1999. (http://www.cmec.ca/stats/pcera/compaper/index.stm)

Yancey, K. B. (1999). Looking back as we look forward: Historicizing writing assessment. *College Composition and Communication, 50,* 483-503.

Yen, W. M. (1998). Investigating the consequential aspects of validity: Who is responsible and what should they do? *Educational Measurement: Issues and Practice, 17,* 5.

Zellermayer, M. (1989). The study of teachers' written feedback to students' writing: Changes in theoretical considerations and the expansion of research context. *Instructional Science, 18,* 145-165

Zinsser, W. (1988). *Writing to learn*. New York: Harper and Row.

APPENDIX 1: WRITING REQUIREMENTS FOR ENGLISH 30-1

General Outcomes	**Specific expectations**
Express ideas and develop understanding	• Draw from a repertoire of effective strategies to form tentative understandings, interpretations and positions • Modify tentative interpretations and tentative positions by weighing and assessing the validity of own and others' ideas, observations and opinions; and identify areas for further inquiry or research
Experiment with language and forms	• Explain how experiments with language, image and structure improve personal craft and increase effectiveness as a text creator. • Experiment with a variety of strategies, activities and resources to explore ideas, observations, opinions, experiences and emotions
Set goals	• Appraise own strengths and weaknesses as a language user and language learner; select appropriate strategies to increase strengths and address weaknesses; monitor the effectiveness of selected strategies; and modify selected strategies as needed to optimize growth
Generate ideas	• Form generalizations and synthesize new ideas by integrating new information with prior knowledge • Draw conclusions that are appropriate to findings, reflect own understandings and are consistent with the identified topic, purpose and situation • Support generalizations and conclusions sufficiently with relevant and consistent detail
Structure texts	• Make revisions as needed to ensure that the beginning of a text in progress establishes purpose and engages audience • Assess the organizational components of a text in progress, and revise them as needed to strengthen their effectiveness as units of thought or experience or to strengthen their contribution to other intended effects • Assess the closing of a text in progress; and revise it as needed to ensure that it is related to purpose, that it establishes a sense of developed understanding and that it will have an appropriate effect on audience • Apply the concepts of unity and coherence to ensure the effective organization of oral, print, visual and multimedia texts • Assess transitions and transitional devices, and revise them as needed to strengthen coherence
Evaluate Sources	• Assess the appropriateness of own and others' understandings and interpretations of works of literature and other texts, by referring to the works and texts for supporting or contradictory evidence
Organize information	• Describe the purpose of inquiry or research and the scope of the inquiry or research topic; identify the target audience; and identify the potential form for the presentation of inquiry or research findings, when applicable • Assess the effectiveness of the controlling idea or desired unifying effect of a text in progress, and refine the controlling idea or desired unifying effect as appropriate to meet the intended purpose • Review the accuracy, specificity, precision, vividness and relevance of details, events, images, facts or other data intended to support a controlling idea or to develop a unifying effect; and add to, modify or delete details, events, images, facts or other data as needed to provide complete and effective support or development • Assess reasoning for logic and evidence for consistency, completeness and relevance; and strengthen reasoning as needed by adding to, modifying or deleting details to provide significant evidence and make effective and convincing arguments • Assess the plausibility and appropriateness of literary interpretations and the precision, completeness and relevance of evidence when reviewing and revising critical/analytical responses to literature
Evaluate information	• Reflect on and describe strategies to evaluate information sources for credibility and bias and for quality; and select, monitor and modify strategies as needed to evaluate sources and detect bias

	• Assess information sources for appropriateness to purpose, audience and presentation form • Assess the accuracy, completeness, currency and relevance of information selected from sources; and assess the appropriateness of the information for purpose. • Identify and describe possible biases and vested interests of sources; and explain how underlying assumptions, biases, and positive or negative spin affect the credibility of sources
Review research process	• Reflect on and describe strategies for developing an inquiry or research plan that will foster understanding, select and monitor appropriate strategies, and modify strategies as needed to plan inquiry or research effectively • Select from a repertoire of effective strategies to develop appropriate inquiry or research plans that will address the topic and satisfy contextual and presentation requirements • Assess the breadth and depth of prior knowledge, and refine questions to further satisfy information needs and to guide the collection of new information • Develop and draw from a repertoire of effective strategies and technologies for gathering, generating and recording information
Expand knowledge of language	• Reflect on personal vocabulary and repertoire of stylistic choices and on their effectiveness; and expand vocabulary and repertoire of stylistic choices • Assess the appropriateness and effectiveness of diction, and revise word choice as needed to create intended effects
Enhance artistry	• Assess the appropriateness and effectiveness of diction, and revise word choice as needed to create intended effects • Explain how stylistic techniques and rhetorical devices are used to create intended effects
Attend to grammar and usage	• Know and be able to apply capitalization and punctuation conventions correctly, including end punctuation, commas, semicolons, colons, apostrophes, quotation marks, hyphens, dashes, ellipses, parentheses, underlining and italics • Know and be able to apply spelling conventions consistently and independently • Understand the importance of grammatical agreement; and assess and revise texts in progress to ensure correctness of grammatical agreement, including correct pronoun reference and pronoun–antecedent agreement, and correct use of modifiers and other parts of speech • Assess and revise texts in progress to ensure correct subject–verb agreement, correct pronoun case and appropriate consistency of verb tense • Use unconventional punctuation, spelling and sentence structure for effect, when appropriate • Assess and revise texts in progress to ensure the correct use of clauses and phrases, including verbal phrases (participle, gerund and infinitive), and to ensure the correct use of structural features • Pay particular attention to punctuation, spelling, grammar, usage and sentence construction when using unfamiliar vocabulary, complex syntax and sophisticated rhetorical devices • Assess strengths and areas of need

APPENDIX 2: ENGLISH 30-1 DIPLOMA EXAM (JUNE 2005)

English Language Arts 30–1 June 2005 Writing Assignments

June 2005

English Language Arts 30–1
Part A: Written Response

Grade 12 Diploma Examination

Description

Time: 2½ hours. This examination was developed to be completed in 2½ hours; however, you may take an additional ½ hour to complete the examination.

Plan your time carefully.

Part A: Written Response contributes 50% of the total English Language Arts 30–1 Diploma Examination mark and consists of two assignments.

- **Personal Response to Texts Assignment**
  Value 20% of total examination mark
- **Critical /Analytical Response to Literary Texts Assignment**
  Value 30% of total examination mark

Recommendation: Read and reflect upon the whole examination before you begin to write. Time spent in planning may result in better writing.

Do not write your name anywhere in this booklet. Feel free to make corrections and revisions directly on your written work.

Instructions

- Complete the Personal Response to Texts Assignment first. The Personal Response to Texts Assignment is designed to allow you time to think and reflect upon the ideas that you may also explore in the Critical /Analytical Response to Literary Texts Assignment.
- Complete **both** assignments.
- You may use the following print references:
  –an English and/or bilingual dictionary
  –a thesaurus
  –an authorized writing handbook
- Space is provided in this booklet for planning and for your written work.
- Use blue or black ink for your written work.

Additional Instructions for Students Using Word Processors

- Format your work using an easy-to-read 12-point or larger font such as Times.
- Double-space your final copy.
- Staple your final printed work to the pages indicated for word-processed work for each assignment. Hand in all work.
- Indicate in the space provided on the back cover that you have attached word-processed pages.

PERSONAL RESPONSE TO TEXTS ASSIGNMENT
Suggested time: approximately 45 to 60 minutes

Carefully read and consider the texts on pages 4 to 6, and then complete the assignment that follows.

The excerpt from the novel *Snow Falling on Cedars* is unavailable for electronic posting. The bibliographic information for this excerpt is: Guterson, David. *Snow Falling on Cedars*. New York: Vintage Contemporaries, 1995, pp. 438–439.

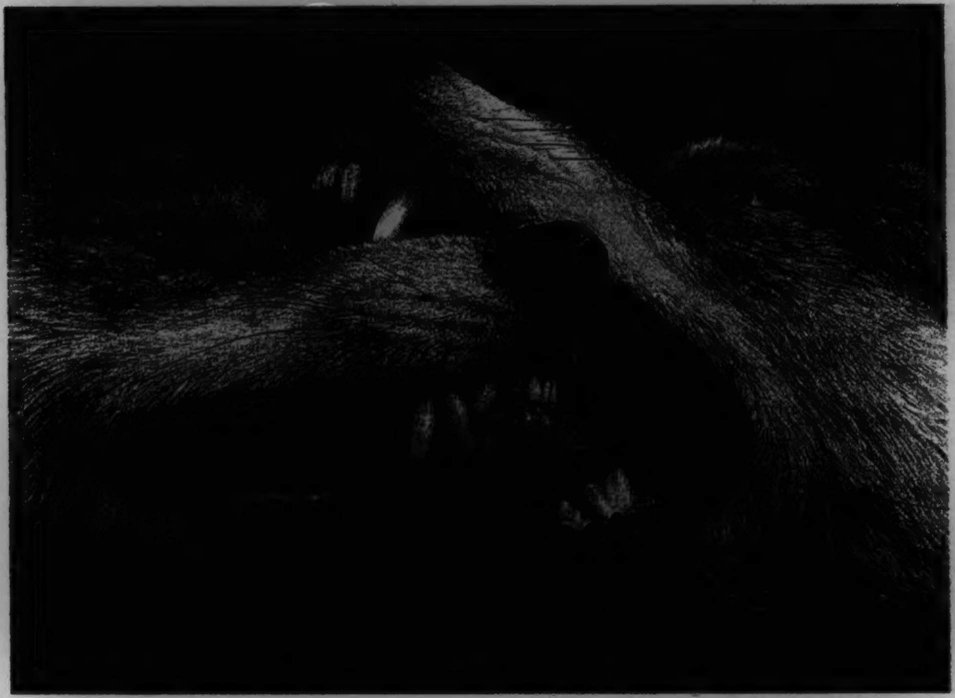
PERSONAL RESPONSE TO TEXTS ASSIGNMENT

The poem "Casting and Gathering" is unavailable for electronic posting. The bibliographic information for this poem is: Heaney, Seamus. "Casting and Gathering." In *Seeing Things*. London: Faber and Faber Limited, 1991, p. 13.

PERSONAL RESPONSE TO TEXTS ASSIGNMENT

This photograph was one of many that Carl Cook included in his print exhibit which was last shown at Seattle's Pacific Science Center in 1988. The photographs for the show were taken at Wolf Haven, a wilderness sanctuary for captive-born wolves.

Wolf Greeting—Rogue & Pretty Girl—1986



When you greet your dog with a playful squeeze on the nose, he usually responds with an outpouring of affection—and maybe a little slobber on your hand. This moment of bonding is important. By displaying dominance (wrapping your hand around his muzzle) and mixing it with affection, you have reassured him of his place within his pack—your family, and this reassurance is vital to your dog's well being. Dominant/submissive behavior is the glue that holds the wolf—and dog pack—together. Deep down inside every dog, from teacup poodle to Great Dane beats the heart of their common ancestor, the wolf.

The greeting you and your dog exchange has the same meaning as the behavior exhibited here with Rogue greeting his mate, Pretty Girl.

Photograph and text by Carl Cook. © Copyright 1986, 1996, Carl Cook. All Rights Reserved.

Fold and tear along perforation.

PERSONAL RESPONSE TO TEXTS ASSIGNMENT
Suggested time: approximately 45 to 60 minutes

The Assignment

In the excerpt from *Snow Falling on Cedars*, the writer describes the dual consequences of the islanders' prolonged mutual dependence. In the poem "Casting and Gathering," the speaker observes the tension that exists between opposing perspectives. "Wolf Greeting—Rogue & Pretty Girl—1986," a photograph of two timber wolves, reflects the interplay between dominant and submissive impulses.

What do these texts suggest to you about the ways in which the desire for independence and the need for security shape an individual's identity? Support your idea(s) with reference to one or more of the texts presented and to your previous knowledge and/or experience.

In your writing, you must

- select a *prose form* that is appropriate to the ideas you wish to express and that will enable you to effectively communicate to the reader
- discuss ideas and/or impressions that are relevant to this assignment

Personal Response to Texts Assignment

Initial Planning

You may respond from a personal, critical, and/or creative perspective. Keep in mind that regardless of the form you choose, you must communicate clearly to the reader.

Briefly identify your choice of prose form, your reason(s) for choosing this prose form, and what you intend to communicate.

Additional space is provided for planning in the examination booklet.

CRITICAL / ANALYTICAL RESPONSE TO LITERARY TEXTS ASSIGNMENT

Suggested time: approximately 1½ to 2 hours

Reflect on the ideas and impressions that you discussed in the Personal Response to Texts Assignment concerning the ways in which the desire for independence and the need for security shape an individual's identity.

The Assignment

Consider how the desire for independence and the need for security have been reflected and developed in a literary text or texts you have studied. Discuss the idea(s) developed by the text creator(s) about an individual's attempt to reconcile the desire to act independently with the need for security.

In your planning and writing, consider the following instructions.

- You must focus your discussion on a literary text or texts *other than* the texts provided in this examination booklet.

- When considering the work(s) that you know well, select a literary text or texts meaningful to you and relevant to this assignment. Choose from short stories, novels, plays, screenplays, poetry, films, or other literary texts that you have studied in English Language Arts 30–1.

- Carefully consider your *controlling idea* or how you will create a strong *unifying effect* in your response.

- As you develop your ideas, support them with appropriate, relevant, and meaningful examples from literary text(s).

Fold and tear along perforation.

CRITICAL / ANALYTICAL RESPONSE TO LITERARY TEXTS ASSIGNMENT

Initial Planning

You may use this space for your initial planning. This information assists markers in identifying the text(s) you have chosen to support your ideas. The markers who read your composition will be very familiar with the literary text(s) you have chosen.

Literary Text(s) and ______________________________
Text Creator(s)
____
____
____

Personal Reflection on Choice of Literary Text(s)
Suggested time: 10 to 15 minutes

Briefly explore your reasons for selecting the literary text(s) you have identified above as support for your response. You may respond in point form, using a diagram such as a mind map, or in another format of your choice. Markers will consider the information you provide here when considering the effectiveness of your supporting evidence.

____
____
____
____
____
____
____
____

Additional space is provided for *Personal Reflection on Choice of Literary Text(s)* in the examination booklet.

Scoring Categories and Scoring Criteria for 2004–2005 Personal Response to Texts Assignment

Ideas and Impressions (10% of total examination mark)
Cross-Reference to the *Program of Studies for*
Senior High School English Language Arts 2.1 2.2 2.3 4.1

When marking **Ideas and Impressions**, the marker should consider the quality of

- the student's **exploration** of the topic
- the student's **ideas and reflection**
- **support** in relation to the student's ideas and impressions

Excellent	**E**	The student's exploration of the topic is insightful. Perceptions and/or ideas are confident and discerning. Support is precise and aptly reinforces the student's ideas and impressions.
Proficient	**Pf**	The student's exploration of the topic is purposeful. Perceptions and/or ideas are thoughtful and considered. Support is specific and strengthens the student's ideas and impressions.
Satisfactory	**S**	The student's exploration of the topic is generalized. Perceptions and/or ideas are straightforward and relevant. Support is adequate and clarifies the student's ideas and impressions.
Limited	**L**	The student's exploration of the topic is vague. Perceptions and/or ideas are superficial and/or ambiguous. Support is imprecise and/or ineffectively related to the student's ideas and impressions.
Poor	**P**	The student's exploration of the topic is minimal. Perceptions and/or ideas are underdeveloped and/or irrelevant. Support is lacking and/or unrelated to the student's ideas and impressions.
Insufficient	**INS**	**Insufficient is a special category. It is not an indicator of quality.** Assign insufficient when • the student has written so little that it is not possible to assess Ideas and Impressions **OR** • the marker can discern no evidence of an attempt to address the task presented in the assignment

Because students' responses to the *Personal Response to Texts Assignment* vary widely—from philosophical discussions to personal narratives to creative approaches—assessment of the *Personal Response to Texts Assignment* on the diploma examination will be in the context of Louise Rosenblatt's suggestion:

> *. . . the evaluation of the answers would be in terms of the amount of evidence that the youngster has actually read something and thought about it, not a question of whether, necessarily, he has thought about it the way an adult would, or given an adult's "correct" answer.*

Rosenblatt, Louise. "The Reader's Contribution in the Literary Experience: Interview with Louise Rosenblatt." By Lionel Wilson. *English Quarterly* 14, no.1 (Spring, 1981): 3–12.

Markers will also consider Grant P. Wiggins' suggestion that we should assess students' writing "with the tact of Socrates: tact to respect the student's ideas enough to enter them fully—even more fully than the thinker sometimes—and thus, the tact to accept apt but unanticipatable or unique responses."

Wiggins, Grant P. *Assessing Student Performance: Exploring the Purpose and Limits of Testing.* San Francisco: Jossey-Bass Publishers, 1993, p. 40.

Scoring Categories and Scoring Criteria for 2004–2005 Personal Response to Texts Assignment (continued)

Presentation (10% of total examination mark)
Cross-Reference to the *Program of Studies for Senior High School English Language Arts* 3.1 3.2 4.1 4.2

When marking Presentation, the marker should consider the effectiveness of

- **voice** in relation to the context created by the student in the chosen prose form
- **stylistic choices** (including quality of language and expression) and the student's creation of **tone**
- the student's development of a **unifying effect**

Consider the proportion of error in terms of the complexity and length of the response.

Excellent **E**	The voice created by the student is convincing. Stylistic choices are precise and the student's creation of tone is adept. The unifying effect is skillfully developed.
Proficient **Pf**	The voice created by the student is distinct. Stylistic choices are specific and the student's creation of tone is competent. The unifying effect is capably developed.
Satisfactory **S**	The voice created by the student is apparent. Stylistic choices are adequate and the student's creation of tone is conventional. The unifying effect is appropriately developed.
Limited **L**	The voice created by the student is indistinct. Stylistic choices are imprecise and the student's creation of tone is inconsistent. The unifying effect is inadequately developed.
Poor **P**	The voice created by the student is obscure. Stylistic choices impede communication and the student's creation of tone is ineffective. A unifying effect is absent.

Scoring Categories and Scoring Criteria for 2004–2005 Critical / Analytical Response to Literary Texts Assignment

Thought and Understanding (7.5% of total examination mark)
Cross-Reference to the *Program of Studies for Senior High School English Language Arts* 2.1 2.2 4.1 4.2

When marking Thought and Understanding, the marker should consider:

- how effectively the student's **ideas** relate to the assignment
- the quality of the **literary interpretations** and **understanding**

Excellent	**E**	Ideas are insightful and carefully considered, demonstrating a comprehension of subtle distinctions in the literary text(s) and the topic. Literary interpretations are perceptive and illuminating.
Proficient	**Pf**	Ideas are thoughtful and considered, demonstrating a competent comprehension of the literary text(s) and the topic. Literary interpretations are revealing and sensible.
Satisfactory	**S**	Ideas are relevant and straightforward, demonstrating a generalized comprehension of the literary text(s) and the topic. Literary interpretations are general but plausible.
Limited	**L**	Ideas are superficial or oversimplified, demonstrating a weak comprehension of the literary text(s) and the topic. Literary interpretations are incomplete and/or literal.
Poor	**P**	Ideas are largely absent or irrelevant, and/or do not develop the topic. Little comprehension of the literary text(s) is demonstrated.
Insufficient	**INS**	**Insufficient is a special category. It is not an indicator of quality.** Assign insufficient when • the student has written so little that it is not possible to assess Thought and Understanding and/or Supporting Evidence **OR** • no reference has been made to literature studied **OR** • the only literary reference present is to the text(s) on the examination **OR** • the marker can discern no evidence of an attempt to address the task presented in the assignment

Because students' responses to the *Critical / Analytical Response to Literary Texts Assignment* vary widely—from philosophical discussions to personal narratives to creative approaches—assessment of the *Critical / Analytical Response to Literary Texts Assignment* on the diploma examination will be in the context of Louise Rosenblatt's suggestion:

> *. . . the evaluation of the answers would be in terms of the amount of evidence that the youngster has actually read something and thought about it, not a question of whether, necessarily, he has thought about it the way an adult would, or given an adult's "correct" answer.*

Rosenblatt, Louise. "The Reader's Contribution in the Literary Experience: Interview with Louise Rosenblatt." By Lionel Wilson. *English Quarterly* 14, no.1 (Spring, 1981): 3–12.

Markers will also consider Grant P. Wiggins' suggestion that we should assess students' writing "with the tact of Socrates: tact to respect the student's ideas enough to enter them fully—even more fully than the thinker sometimes—and thus, the tact to accept apt but unanticipatable or unique responses."

Wiggins, Grant P. *Assessing Student Performance: Exploring the Purpose and Limits of Testing.* San Francisco: Jossey-Bass Publishers, 1993, p. 40.

Scoring Categories and Scoring Criteria for 2004–2005 Critical / Analytical Response to Literary Texts Assignment (continued)

Supporting Evidence (7.5% of total examination mark)
Cross-Reference to the *Program of Studies for*
Senior High School English Language Arts 2.3 3.2 4.1 4.2

When marking Supporting Evidence, the marker should consider

- **the selection and quality of evidence**
- **how well the supporting evidence is integrated, synthesized, and/or developed to support the student's ideas**

Consider ideas presented in the *Personal Reflection on Choice of Literary Text(s).*

Excellent	**E**	Support is explicit, precise, and deliberately chosen to reinforce the student's ideas in an effective and judicious way. A strong connection to the student's ideas is maintained.
Proficient	**Pf**	Support is relevant, accurate, and occasionally deliberately chosen to reinforce the student's ideas in a logical and clear way. A clear connection to the student's ideas is maintained.
Satisfactory	**S**	Support is appropriate, general, and adequately reinforces the student's ideas but occasionally may lack persuasiveness and/or consistency. A straightforward connection to the student's ideas is maintained.
Limited	**L**	Support is repetitive, contradictory, and/or ambiguous, and may be inappropriate or merely a restatement of what was read. The connection to the student's ideas is vague and/or unclear.
Poor	**P**	Support is irrelevant, overgeneralized, and/or lacking. The support, if present, is largely unrelated to any idea(s) that may be present.

Scoring Categories and Scoring Criteria for 2004–2005 Critical/Analytical Response to Literary Texts Assignment (continued)

Form and Structure (5% of total examination mark)
Cross-Reference to the *Program of Studies for Senior High School English Language Arts* 2.2 3.1 4.1 4.2

When marking **Form and Structure**, the marker should consider how effectively the student's organizational choices result in

- a **coherent, focused, and shaped** discussion in response to the assignment
- a **unifying effect** or a **controlling idea** that is developed and maintained

Excellent **E**	An effective arrangement of ideas and details contributes to a fluent and shaped discussion that is developed skillfully. The unifying effect or controlling idea is consistently sustained and integrated.
Proficient **Pf**	A purposeful arrangement of ideas and details contributes to a competent and controlled discussion that is developed capably. The unifying effect or controlling idea is coherently sustained and presented.
Satisfactory **S**	A straightforward arrangement of ideas and details provides direction for the discussion that is developed appropriately. The unifying effect or controlling idea is presented and maintained generally; however, coherence may falter.
Limited **L**	A discernible but ineffectual arrangement of ideas and details provides some direction for the discussion that is not deliberately developed. A unifying effect or controlling idea is inconsistently maintained.
Poor **P**	A haphazard arrangement of ideas and details provides little or no direction for the discussion, and development is lacking or obscure. A unifying effect or controlling idea is absent.

Scoring Categories and Scoring Criteria for 2004–2005 Critical / Analytical Response to Literary Texts Assignment (continued)

Matters of Choice (5% of total examination mark)
Cross-Reference to the *Program of Studies for Senior High School English Language Arts* 4.2

When marking **Matters of Choice**, the marker should consider how effectively the student's choices enhance communication. The marker should consider

- **diction**
- choices of **syntactic structures** (such as parallelism, balance, inversion)
- the extent to which **stylistic choices** contribute to the creation of voice

Excellent **E**	Diction is precise. Syntactic structures are effective and sometimes polished. Stylistic choices contribute to a confident composition with a convincing voice.
Proficient **Pf**	Diction is specific. Syntactic structures are generally effective. Stylistic choices contribute to a competent composition with a capable voice.
Satisfactory **S**	Diction is adequate. Syntactic structures are straightforward, but attempts at complex structures may be awkward. Stylistic choices contribute to a clear composition with an appropriate voice.
Limited **L**	Diction is imprecise and/or inappropriate. Syntactic structures are frequently awkward and/or ambiguous. Inadequate language choices contribute to a vague composition with an undiscerning voice.
Poor **P**	Diction is overgeneralized and/or inaccurate. Syntactic structures are uncontrolled and/or unintelligible. A lack of language choices contributes to a confused composition with an ineffective voice.

Scoring Categories and Scoring Criteria for 2004–2005 Critical/Analytical Response to Literary Texts Assignment (continued)

Matters of Correctness (5% of total examination mark)
Cross-Reference to the *Program of Studies for Senior High School English Language Arts* 4.2

When marking **Matters of Correctness**, the marker should consider the correctness of …

- **sentence construction** (completeness, consistency, subordination, coordination, predication)
- **usage** (accurate use of words according to convention and meaning)
- **grammar** (subject-verb/pronoun-antecedent agreement, pronoun reference, consistency of tense)
- **mechanics** (punctuation, spelling, capitalization)

Consider the proportion of error in terms of the complexity and length of the response.

Excellent **E**	This writing demonstrates confidence in control of correct sentence construction, usage, grammar, and mechanics. The relative absence of error is impressive considering the complexity of the response and the circumstances.
Proficient **Pf**	This writing demonstrates competence in control of correct sentence construction, usage, grammar, and mechanics. Minor errors in complex language structures are understandable considering the circumstances.
Satisfactory **S**	This writing demonstrates control of the basics of correct sentence construction, usage, grammar, and mechanics. There may be occasional lapses in control and minor errors. However, the communication remains clear.
Limited **L**	This writing demonstrates faltering control of correct sentence construction, usage, grammar, and mechanics. The range of errors blurs the clarity of communication.
Poor **P**	This writing demonstrates lack of control of correct sentence construction, usage, grammar, and mechanics. Jarring errors impair communication.

APPENDIX 3: STUDENT SURVEY

Teaching Writing in English 30-1 Participant Number__________
Student Survey

Once you have completed this survey, please place it in a sealed envelope and return it to your English 30-1 teacher.

This survey is designed to explore your views of what it means to write effectively. It is also designed to explore your perceptions of the writing component of the English 30-1 diploma exam and its views regarding what it means to write effectively.

PART I) STUDENT BACKGROUND

For the following questions, please circle the response that best matches your background.

1. What is your gender? **Male Female**
2. Is English the language that is spoken in your home? **yes no**
3. Have you ever written a practice English 30-1 diploma exam (writing component)? **yes no**
4. Is your English 30-1 course semestered? **yes no**
5. In percentages, what final grade do you expect to receive for the **classroom-based** portion of English 30-1? __________
6. In percentages, what grade do you expect to receive for the **exam-based** portion of English 30-1? __________
7. How would you rate your writing ability? **excellent above average average below average**
8. How would you rate your overall ability as a student? **excellent above average average below average**

PART II) BELIEFS AND ATTITUDES

The following set of questions explores your thoughts about writing, teaching, and testing. Using the scale below, please indicate your level of agreement with the following statements by selecting the appropriate number.

strongly agree = 5, agree = 4, neutral = 3, disagree =2, strongly disagree = 1

		SA	A	N	D	SD
9.	I believe that an effective writing process is necessary for excellent performance on the exam.	5	4	3	2	1
10.	I believe that reshaping is a very important element of the writing process.	5	4	3	2	1
11.	Studying test taking skills will improve my performance on the exam.	5	4	3	2	1
12.	If my teacher concentrates his/her instruction on things covered by the diploma exam, I will score better on the exam.	5	4	3	2	1
13.	If my teacher concentrates his/her instruction on things covered by the diploma exam, my overall learning will improve.	5	4	3	2	1
14.	I believe that an effective writing process is the foundation of good writing.	5	4	3	2	1
15.	I think first draft writing should be effectively organized.	5	4	3	2	1
16.	For me, clarity of expression is an important goal of first draft writing.	5	4	3	2	1
17.	For me, clarity of ideas is an important goal in first draft writing.	5	4	3	2	1
18.	I write best on assignments that are meaningful to me personally.	5	4	3	2	1
19.	My English 30-1 teacher focuses his/her instruction on preparing us for the writing part of the English 30-1 diploma exam	5	4	3	2	1

PART III) CLASSROOM ACTIVITIES

The following set of questions explores your writing habits, and the types of activities you complete in your English 30-1 class.

Using the scale below, please indicate your level of agreement with the following statements by circling the number that best represents your opinion.

always = 5, often = 4, sometimes = 3, seldom = 2, never = 1

	When studying writing in my English 30-1 class,	A	O	So	Se	N
20.	In English 30-1, I am expected to write using the five paragraph essay.	5	4	3	2	1
21.	In English 30-1, I have completed grammar and/or punctuation worksheets.	5	4	3	2	1
22.	In English 30-1, I personally conference with my teacher about my writing.	5	4	3	2	1
23.	In English 30-1, I am given class time to conference with my classmates about their writing.	5	4	3	2	1
24.	In English 30-1, I mainly receive feedback on my writing through written comments.	5	4	3	2	1
25.	In English 30-1, I am taught about writing through lectures	5	4	3	2	1
26.	In English 30-1, I am taught about writing through group projects	5	4	3	2	1
27.	In English 30-1, my teacher uses practice diploma exam questions when teaching writing.	5	4	3	2	1
28.	In English 30-1, I begin writing assignments the day before they are due	5	4	3	2	1
29.	In English 30-1, I complete writing assignments using multiple-drafts	5	4	3	2	1
30.	In English 30-1, I try to get my first draft good enough so that I don't have to edit or revise it later	5	4	3	2	1
31.	In English 30-1, I organize my ideas before I start writing	5	4	3	2	1
32.	Whenever possible I write using multiple drafts	5	4	3	2	1
33.	In English 30-1, I complete writing assignments the day before they are due	5	4	3	2	1

PART IV) WRITING SKILLS

The following set of questions deals with skills used when writing. For each item you are asked to do two things. Using the lefthand column: "Classroom assignments", indicate how important each skill is for getting good grades on your **English 30-1 classroom writing assignments**. Using the right hand column: "Exam", indicate how important you think each skill is for doing well on **the writing component of the English 30-1 diploma exam**.

Importance for classroom assignments

very important = 5,
important = 4,
somewhat important = 3,
minimally important = 2,
not important at all = 1

Importance for exam performance

very important = 5,
important = 4,
somewhat important = 3,
minimally important = 2,
not important at all = 1

Example: How important are the following skills for successfully completing and receiving high marks on your classroom writing assignments and on your diploma exam questions? Your ability to:

Importance for class assignments		**Importance for Diploma exam**
5 4 3 2 1	1. Penmanship	5 4 3 2 1

If writing legibly is somewhat important to your teacher's marking but not important at all for people scoring the diploma exam you would circle a 3 under the column: **Importance for classroom assignment**, and a 1 under the column: **Importance for diploma exam**

Importance for class assignments		**Importance for Diploma exam**
5 4 [✓ 3] 2 1	1. penmanship	5 4 3 2 [✓ 1]

Importance for classroom assignments

very important = 5,
important = 4,
somewhat important = 3,
minimally important = 2,
not important at all = 1

Importance for exam performance

very important = 5,
important = 4,
somewhat important = 3,
minimally important = 2,
not important at all = 1

In English 30-1 how important are the following skills for successfully completing and receiving high marks on your classroom writing assignments and on your diploma exam questions? Your ability to:

Importance for class assignments			**Importance for diploma exam**
5 4 3 2 1	34.	develop a strong voice	5 4 3 2 1
5 4 3 2 1	35.	restructure a first draft	5 4 3 2 1
5 4 3 2 1	36.	develop supporting details	5 4 3 2 1
5 4 3 2 1	37.	generate a first draft with a minimum of errors	5 4 3 2 1
5 4 3 2 1	38.	engage in a multiple-draft writing process	5 4 3 2 1
5 4 3 2 1	39.	employ word-choice purposefully	5 4 3 2 1
5 4 3 2 1	40.	Critically reexamine a first draft	5 4 3 2 1
5 4 3 2 1	41.	use grammar with a minimum of errors	5 4 3 2 1
5 4 3 2 1	42.	employ mechanics (punctuation, spelling, capitalization) with a minimum of errors	5 4 3 2 1
5 4 3 2 1	43.	generate ideas quickly	5 4 3 2 1
5 4 3 2 1	44.	develop coherent, focused, and shaped ideas	5 4 3 2 1
5 4 3 2 1	45.	use a multiple-draft writing process effectively	5 4 3 2 1
5 4 3 2 1	46.	construct sentences properly	5 4 3 2 1
5 4 3 2 1	47.	write effectively under pressure	5 4 3 2 1
5 4 3 2 1	48.	develop literary interpretations	5 4 3 2 1

Thank you for completing this survey.